



Lietuvių kalbos teksto sintaksinės-semantinės analizės informacinė sistema

LKSSAIS vystymas

Automatinio dokumentų santraukų sudarymo IT sprendimo naudotojo instrukcija

TURINYS

| | |
|---|----------|
| 1. Įžanga | 3 |
| 1.1. Dokumento paskirtis..... | 3 |
| 1.2. Sutrumpinimai | 3 |
| 1.3. Sutartiniai žymėjimai..... | 3 |
| 1.4. Panaudotų dokumentų sąrašas..... | 3 |
| 1.5. ADSS IT sprendimo paskirtis ir tikslai..... | 3 |
| 2. Bendroji informacija | 5 |
| 2.1. Įvedimo ir valdymo elementai | 5 |
| 2.2. Pagrindinis langas | 5 |
| 3. ADSS naudotojo instrukcija | 6 |
| 3.1. Santraukos sudarymas..... | 6 |
| 3.2. Santraukos peržiūra | 7 |
| 3.3. Santraukos pašalinimas..... | 8 |
| 3.4. Nustatymų keitimas | 8 |
| 3.5. Projekto Semantika-2 komponentų nustatymai | 8 |

1. Įžanga

1.1. Dokumento paskirtis

Šio dokumento paskirtis – pateikti modernizuojamos LKSSAIS automatinio dokumentų santraukų sudarymo (ADSS) IT sprendimo naudotojo instrukcijas.

1.2. Sutrumpinimai

1.1 lentelė. Sutrumpinimai

| Santrumpa | Paaškinimas |
|----------------|--|
| LKSSAIS | Lietuvių kalbos teksto sintaksinės-semantinės analizės informacinė sistema |
| ADSS | Automatinis dokumentų santraukų sudarymas |
| VDU vykdytojai | Užsakovo (VDU) projekto vykdytojų komanda |
| Diegėjas | Programavimo ir diegimo paslaugų konkursą laimėjusi įmonė (UAB ATEA) |

1.3. Sutartiniai žymėjimai

1.2 lentelė. Sutartiniai žymėjimai

| Žymėjimas | Aprašymas |
|-------------------|--|
| Laukas | Duomenų įvedimo laukas, matomas aprašomame lange. Žymimas paryškintu tekstu. |
| [Mygtukas] | Dialogo lango arba lango mygtukai (taip pat vadinami komandiniais mygtukais) tekste yra žymimi kvadratiniais skliausteliais []. Tarp skliaustelių yra rašomas mygtuko pavadinimas. Sistemoje atvaizduoti kaip keturkampiai mygtukai. |
| [Nuoroda] | Dialogo lango arba lango nuorodos. |
| Kortelė | Išsikišusi lango dalis, kurią spustelėjus atveriamas atitinkamas langas. |

1.4. Panaudotų dokumentų sąrašas

1.3 lentelė. Panaudotų dokumentų sąrašas

| Eil. Nr. | Pavadinimas |
|----------|---|
| 1. | ADSS IT sprendimo detalios analizės ir projektavimo specifikacija |

1.5. ADSS IT sprendimo paskirtis ir tikslai

Čia aprašomas ADSS IT sprendimas yra skirtas automatiškai sudaryti lietuvių kalbos tekstų santraukas bei nustatyti teksto raktažodžius. ADSS IT sprendimas skirtas žiniasklaidos, administraciniams ar teisės tekstams. Šis IT sprendimas naudodamasis mašininio apmokymo algoritmais nustato raktažodžius ir sakinius, kurie geriausiai iliustruoja duotojo teksto turinį. Jo įeigoje turi būti lietuvių kalbos, UTF-8 koduotės tekstas.

ADSS IT sprendimas naudoja projekte modernizuotas interneto tarnybas. Anotatorius: segmentavimo komponento interneto tarnybą (LEX), morfologijos interneto tarnybą kalbos dalių

ir antraštinių žodžių formų (lemų) nustatymui (MORPH), įvardytųjų esybių nustatymo komponento tarnybą (NER) ir santraukų nustatymo komponentą (ADSS).

Santraukas galima sudaryti visiems failams, kuriuos palaiko *Apache Tika* įrankis¹.

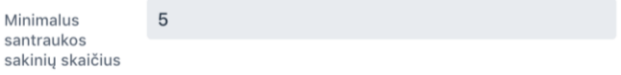
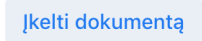


¹ <https://tika.apache.org/1.24.1/formats.html>

2. Bendroji informacija

Šiame skyriuje pateikiama informacija apie įvedimo ir valdymo elementus, naudojamus sistemoje bei bendrąsias sistemos funkcijas.

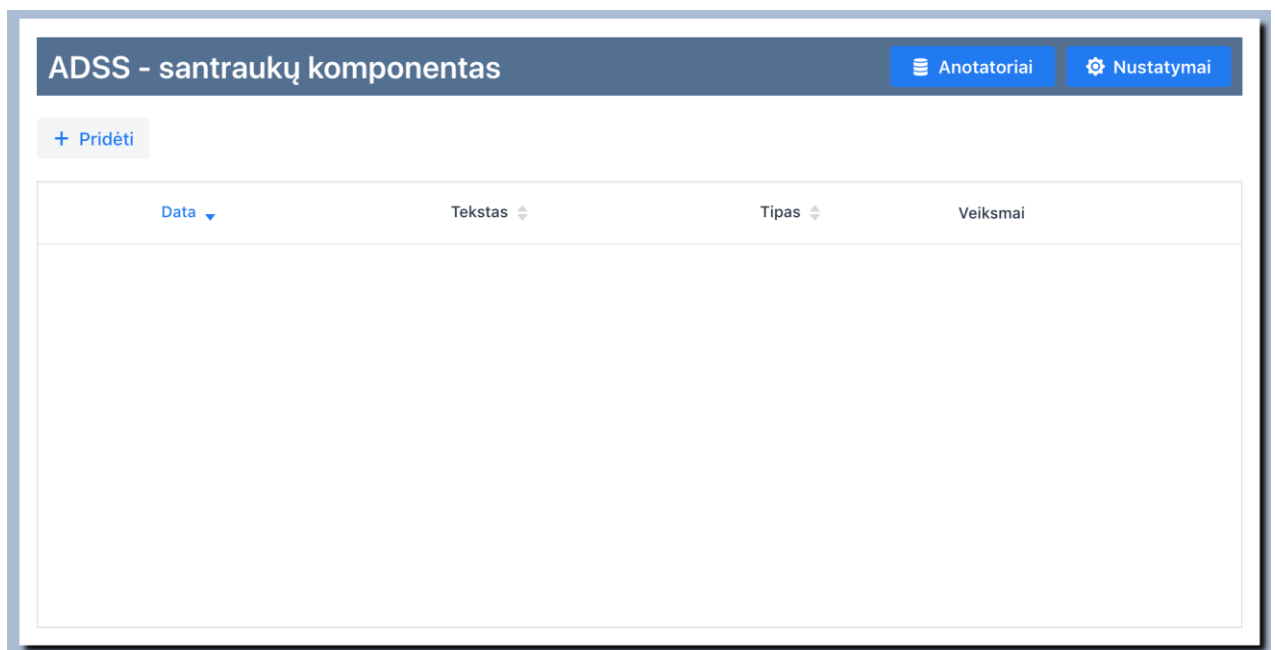
2.1. Įvedimo ir valdymo elementai

2.1 lentelė. Įvedimo ir valdymo elementai

| Elementas | Aprašymas |
|---|--|
|  | Įvedimo elementas. Šiame įvedimo elemente priklausomai nuo lauko paskirties galima įvesti arba skaičiaus reikšmės. |
|  | Failo pasirinkimo elementas. Paspaudus mygtuką [Įkelti dokumentą] atveriamas standartinis failo pasirinkimo langas. |
|  | Žymimasis langelis (angl. <i>checkbox</i>) skirtas „Taip“ – „Ne“ reikšmių nustatymo elementas. Galimos reikšmės: pažymėtas (reikšmė „Taip“) arba nepažymėtas (reikšmė „Ne“). |
|  | Parametrų jungiklis (angl. <i>radio button</i>) skirtas pasirinkti vieną parametą iš trijų. Aprašymas atspindi galimą veiksmą ar pasirinkimą. Pažymėtas parametrų jungiklis parodo, kuris vienas iš kelių galimų veiksmų yra pasirinktas. Sistemos naudotojas gali pasirinkti norimą parametrų jungiklį pelės paspaudimu. |

2.2. Pagrindinis langas

Svarbiausias navigacijos komponentas yra pagrindinis langas (2.1 pav. Pagrindinis sistemos langas). Šiame lange atliekami visi pagrindiniai veiksmai.

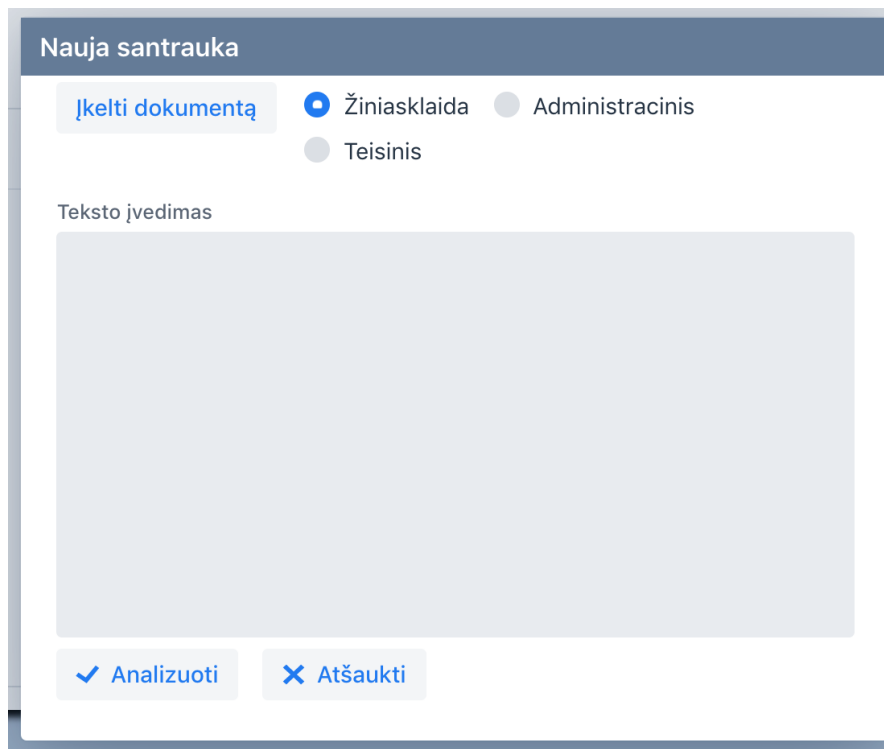


2.1 pav. Pagrindinis sistemos langas

3. ADSS naudotojo instrukcija

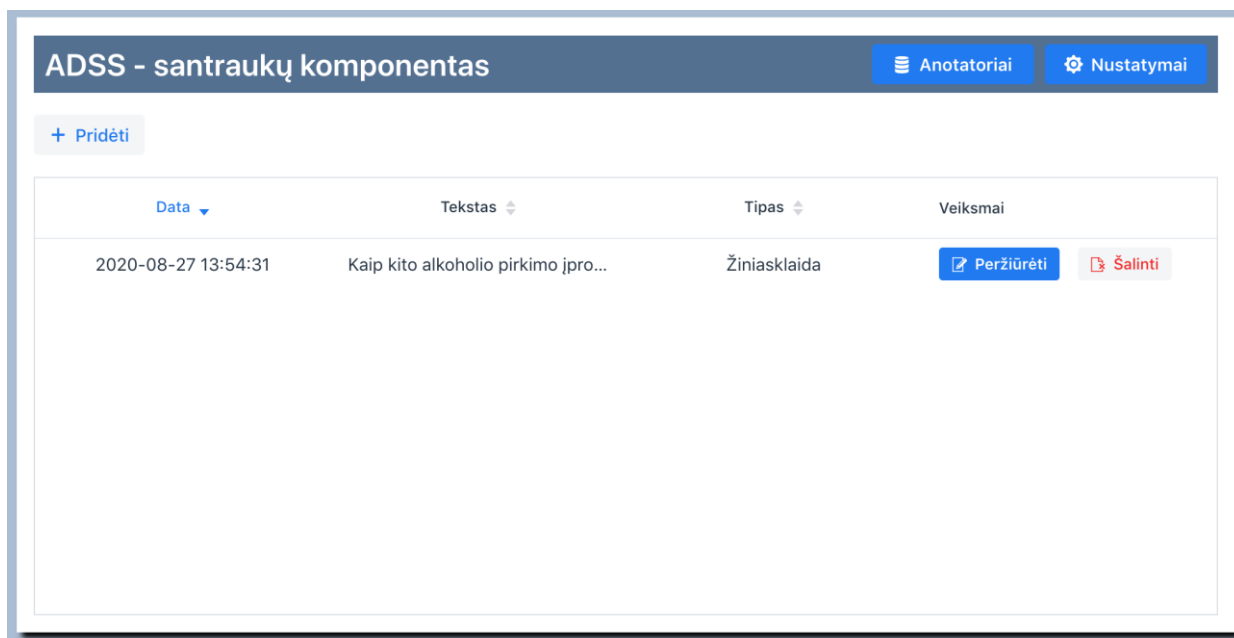
3.1. Santraukos sudarymas

1. Naudotojas per naršyklę atveria pagrindinį sistemos langą.
2. Naudotojas paspaudžia mygtuką **[+Pridėti]**. Atsidaro langas (3.1 pav. Nauja santrauka langas)



3.1 pav. Nauja santrauka langas

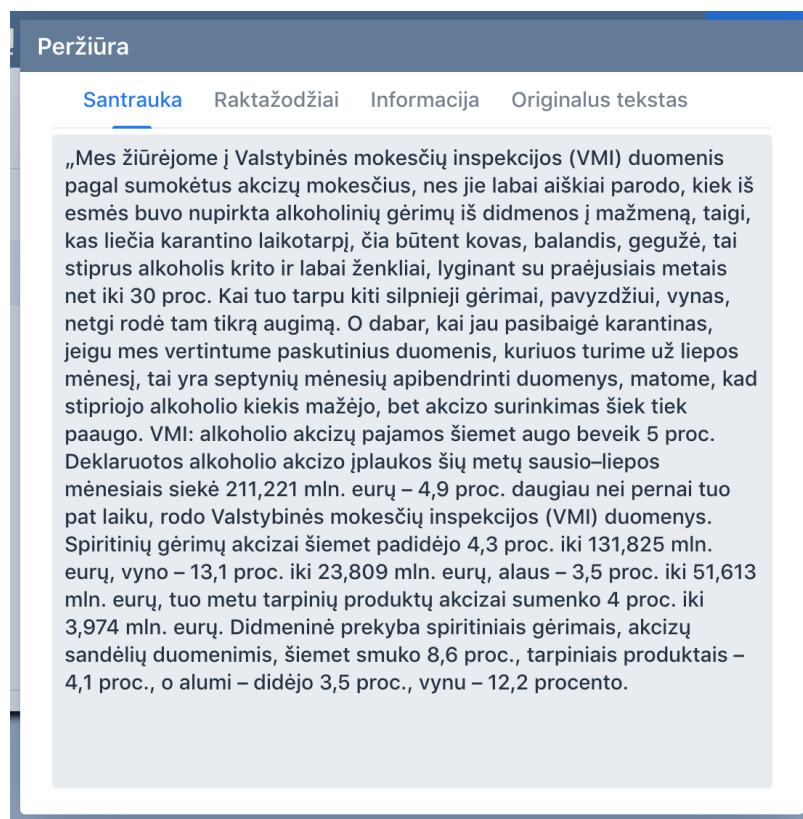
3. Naudotojas gali pasirinkti ar sudaryti santrauką iš failo ar įklijuoti tekstą:
 - a. Naudotojas įklijuoja tekstą tiesiai į **teksto įvedimo** lauką
 - b. Naudotojas pasirenka įkelti failą. Naudotojui paspaudus mygtuką **[Įkelti dokumentą]**, atveriamas lokalaus kompiuterio failų pasirinkimo langas, kuriame naudotojas pasirenka norimą failą. Lauką užpildyti privaloma. Pasirinktas failo turinys yra nuskaitomas ir atvaizduojamas **teksto įvedimo** lauke.
4. Naudotojas pasirenka dokumento tipą (žiniasklaida, administracinis ar teisė)
5. Naudotojas paspaudžia mygtuką **[Analizuoti]**.
6. Pateikiamas pranešimas „Anotuojama“ ir „Santrauka sukurta“.
7. Pagrindiniame lange atsiranda naujas santraukos įrašas (3.2 pav. Naujas santraukos įrašas).



3.2 pav. Naujas santraukos įrašas

3.2. Santraukos peržiūra

1. Naudotojas per naršyklę atveria pagrindinį sistemos langą.
2. Naudotojas pasirenka santrauką, kurią nori peržiūrėti.
3. Naudotojas paspaudžia mygtuką **[Peržiūrėti]**.
4. Atsidaro santraukos peržiūros langas (3.3 pav. Santraukos peržiūros langas).



3.3 pav. Santraukos peržiūros langas

5. Naudotojas gali peržiūrėti:
 - a. Sudarytą teksto santrauką – mygtukas **[Pateikti]**.
 - b. Pagrindinius teksto raktažodžius – mygtukas **[Raktažodžiai]**.
 - c. Struktūrinę teksto informaciją – mygtukas **[Informacija]**.
 - d. Originalų tekstą – mygtukas **[Originalus tekstas]**.

3.3. Santraukos pašalinimas

1. Naudotojas per naršyklę atveria pagrindinį sistemos langą.
2. Naudotojas pasirenka santrauką, kurią nori pašalinti.
3. Naudotojas paspaudžia mygtuką **[Šalinti]**.
4. Pateikiamas pranešimas „Sėkmingai pašalinta“.
5. Santrauka yra pašalinta.

3.4. Nustatymų keitimas

1. Naudotojas per naršyklę atveria pagrindinį sistemos langą.
2. Naudotojas paspaudžia mygtuką **[Nustatymai]**.
3. Atsidaro nustatymų peržiūros langas (3.4 pav.).

| Nustatymai | |
|--|--------------------------|
| Minimalus santraukos sakinių skaičius | 3 |
| Maksimalus santraukos sakinių skaičius | 10 |
| Maksimalus raktažodžių skaičius | 10 |
| Išplėstinė informacija | <input type="checkbox"/> |
| <div>✓ Saugoti ✗ Atšaukti ↻ Atstatyti</div> | |

3.4 pav. Nustatymų langas

4. Naudotojas gali pakeisti šiuos parametrus:
 - a. Minimalų santraukos sakinių skaičių.
 - b. Maksimalų santraukos sakinių skaičių.
 - c. Maksimalų pagrindinių raktažodžių skaičių.
 - d. Uždėjus varnelę ant **Išplėstinė informacija**, bus rodomi sakinių ir raktažodžių įverčiai.

3.5. Projekto Semantika-2 komponentų nustatymai

1. Naudotojas per naršyklę atveria pagrindinį sistemos langą.
2. Naudotojas paspaudžia mygtuką **[Anotatoriai]**.

3. Atsidaro Projekto *Semantika-2* komponentų nustatymo langas (3.5 pav.).

Anotavimo komponentų nustatymai

| | |
|--------------------------------|---|
| LEX komponento adresas | <code>http://lex-svc.tetragrama.svc.cluster.local:7000</code> |
| MORPH komponento adresas | <code>http://morph-api-svc.fotonija.svc.cluster.local:8090/m</code> |
| NER komponento adresas | <code>http://ner-svc.tetragrama.svc.cluster.local:7002/ner</code> |
| ADSS komponento adresas | <code>http://summary-svc.vdu.svc.cluster.local:5000/adss</code> |

✓ Saugoti

✗ Atšaukti

3.5 pav. Anotatorių nustatymai

4. Naudotojas gali pakeisti projekto *Semantika-2* parsisiųstų ir naudotojo kompiuteryje įdiegtų komponentų adresus:
- LEX* komponento.
 - MORPH* komponento.
 - NER* komponento.
 - ADSS* komponento