

Dabiskas valodas apstrādes metodes un rīki virtuālo asistentu izstrādē

Autors: Signe Bērziņa, sb18068
Darba vadītāja: Dr. Dat., LU profesore Inguna Skadiņa



Darba mērķis

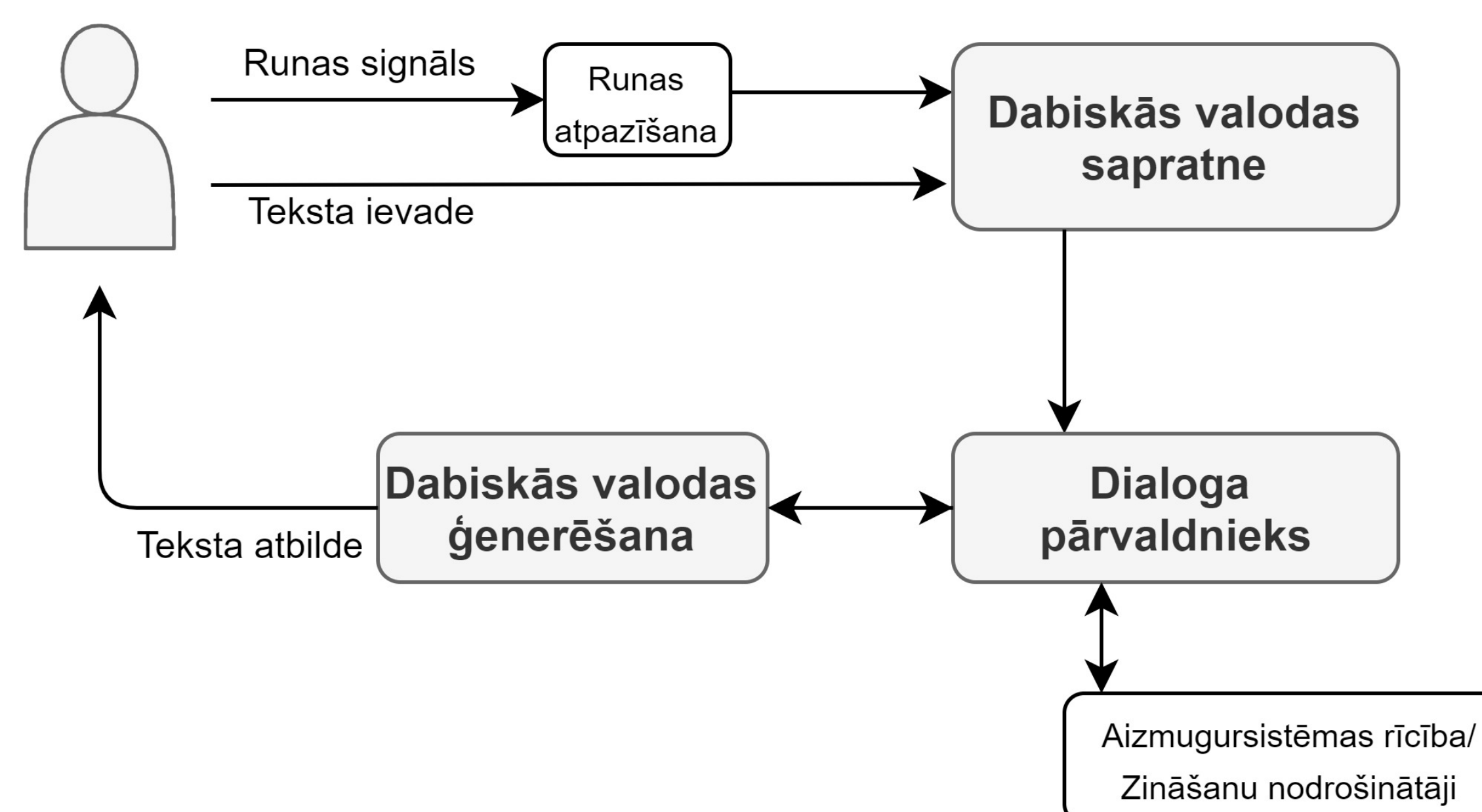
Darba mērķis bija izpētīt pieejamo literatūru par virtuālajiem asistentiem. Aplūkot to veidus un darbības principus, kā arī kādi statistikas modeļi tiek lietoti būvei un kādi rīki ir pieejami.

Dialogu sistēmas

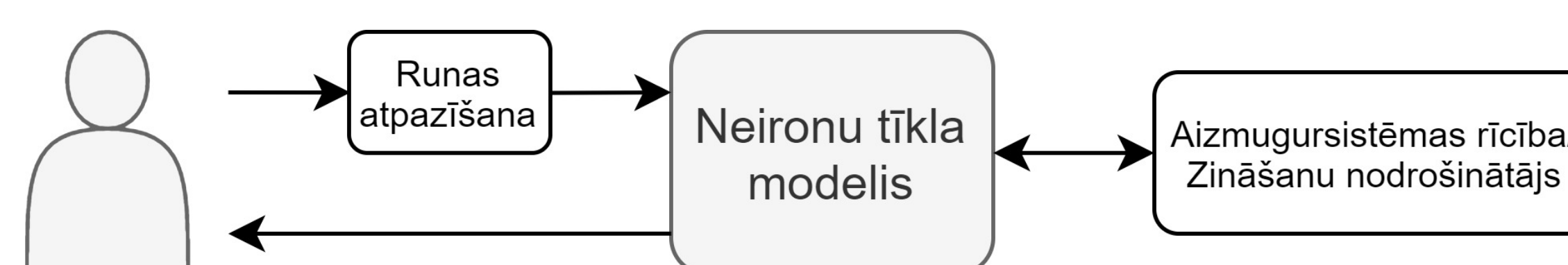
Dialogu sistēmas ir datorprogrammas, kas komunicē ar lietotāju izmantojot dabisko valodu. Tās ir sastopamas salīdzinoši bieži un kļūst arvien populārākas, kā arī tās var sastapt tīmekļa vietnēs, telefonos un citur. Latvijā arī ir daudz dažādu piemēru šādām sistēmām, piemēram, plašais klāsts Hugo un Tildes virtuālo asistentu (*Una - Uznēmumu reģistrs* vai iepriekš pieminēties *Covid-19 virtuālais asistents*).

- **Tērzēšanas sistēmas** - vienkāršākās dialogu sistēmas, kas var turpināt izvērstas sarunas ar mērķi atdarināt neformālu cilvēku komunikāciju raksturīgu nestrukturētai sarunai vai tērzēšanai.
- **Dialogu sistēmas konkrētu uzdevumu veikšanai** - sistēmas, kas izmanto sarunu ar lietotājiem, lai palīdzētu paveikt lietotāja pieprasītos uzdevumus. Tās var paveikt dažādus uzdevumus, piemēram, ziņot par laika apstākļiem un atskaņot dziesmas. Uzbūvēt sarežģītu uzdevumu veikšanas dialogsistēmu joprojām viena no pētāmākajām problēmām.
- **Jautājumu atbilžu dialogu sistēma** - sistēma, kas sniedz īsas un tiešas atbildes uz lietotāja vaicājumu, pamatojoties uz bagātīgām zināšanām, kas iegūtas no dažādiem datu avotiem, tostarp teksta kolekcijām, piemēram, interneta dokumentiem un iepriekš apkopotām zināšanu bāzēm, piemēram, pārdošanas un marketinga datu kopām.

Mijiedarbību starp dialogu sistēmas pamatelementiem



Pilnība datu vadīta dialogu sistēma



Ar neironu tīklu attīstību arvien vairāk un vairāk tiek mēģināts veidot dialogsistēmas nepalaujoties uz iepriekš minētajiem pamatelementiem. Šādas sistēmas ir **pilnībā datu vadīta dialogu sistēma**.

Rīki

- IBM Watson NLP
- Deeppavlov
- Rasa
- Bot Framework SDK
- Amazon Alexa
- wit.ai
- AllenNLP
- Dialogflow
- OpenDialog
- Stanford Almond

Dialogu sistēmas pamatelementi

- **Dabiskās valodas sapratne** - sistēmas mēģinājums izprast lietotāja ievadīto tekstu. Lietotāja izteikums tiek pārvērsts semantiskā attēlojumā, ko var izmantot dialoga pārvaldnieks. Tas ietver morfoloģijas, sintakses un semantikas rīku izmantošanu.
- **Dialoga pārvaldnieks** - pārvalda visus dialoga aspektus, kā piemēram, uztur dialoga vēsturi, pielāgo dialogu, pārvalda iniciatīvu un sistēmas reakciju, kā arī citus aspektus.
- **Atbildes ģenerators** - sistēmas daļā atbildīga par dialoga pārvaldnieka izvēlēti saziņas mērķa pārvēršanu dabiskās valodas formā. Tas ir svarīgs komponents, kas ietekmē dialogu sistēmas dabiskumu un tādējādi arī lietotāja pieredzi.

Dzilā mašīnmācīšanās

Tradicionālajām dialogsistēmām ir sarežģītas. Līdz ar dziļās mašīnmācīšanās sasniegumiem, neironu tīklu modeļi ir ieguvuši popularitāti dialogu sistēmu veidošanā. Neironu tīklus izmanto katrā dialogu sistēmas elementā, piemēram, *dabiskās valodas sapratnē*, jomas noteikšanas un nodoma noteikšanai uzdevuma veikšanai, lietojot neironu tīklus klasifikācijai, tie ir izrādījušies efektīvāki, nekā tradicionālie modeļi.

Secinājumi

Virtuālo dialogu sistēmu būvēšana ir daudz dažādas metodes, ka arī modeļi un rīki ko pielietot. Tika apskatīti dialogu sistēmas veidi un pamatelementi, ka arī sniegs neliels ieskats dialogu sistēmas izveides rīkos. Tas kopā veido izpratni par to, kas ir nepieciešams veidojot dialogu sistēmas.