

Kalbos technologijų pakomisės prioritetizuotų projektinių pasiūlymų pristatymas

Daiva Vaišnienė
Valstybinė lietuvių kalbos komisija

Automatinio vertimo sistemų tobulinimas

- Pritaikymo galimybių išplėtimas (specializuotose srityse, pvz., medicinos, viešojo administravimo).
- Efektyviau integruotas semantinis ir sintaksinis analizatorius (statistinis kalbos modelis).
- Naujų kalbų porų įtraukimas (rusų, lenkų ar vokiečių).
- Naujų aktualių kompiuterinių programų lokalizavimas.

Šnekamosios kalbos technologijų tobulinimas

- Atpažintuvo išplėtimas iki 3000 žodžių.
- Garsyno išplėtimas iki 300 val.
- Keliakalbių šnekos atpažinimo ir sintezavimo įrankių sukūrimas (susiejimas su automatinio vertimo sistemomis).
- Įrankių susiejimas su semantiniu bei sintaksiniu analizatoriumi.

Infrastruktūros plėtra

- Portalo www.raštija.lt funkcionalumo gerinimas.
- Didesnis išteklių ir įrankių integravimas per bendrą paieškos sistemą.
- Informacijos apie kalbos technologijų ir išteklių panaudojimo galimybes sklaida visuomenei ir verslui.

Bazinių kalbos išteklių ir įrankių kūrimas

- Statistinio kalbos modelio (kompiuterinės gramatikos) tobulinimas.
- Anotuoto sintaksinio tekstyno (*treebank*) kūrimas.
- Tezauro (*Wordnet*) kūrimas.

Sintaksinės ir semantinės analizės tobulinimas

- Paieškos įrankių susiejimas su daugiakalbe aplinka (paieška įvairiakalbiuose šaltiniuose, vertimo sistemos integravimas).
- Pritaikymas socialinių medijų, nuomonių analizei (informacijos rinkimas iš socialinių tinklų, saugumo ir kitų grėsmių nustatymas ir pan.).
- Paieškos sistemos išplėtimas ir tobulinimas.
- Klausimų ir atsakymų (dialogo) sistemos kūrimas.
- Santraukų kūrimo įrankiai.
- Semantinės žinių tvarkymo sistemos, skirtos viešojo ir privataus sektoriaus informacijai valdyti.

Prioritetinė tvarka

I

- Automatinio vertimo sistemų tobulinimas

II

- Šnekamosios kalbos technologijų tobulinimas

III

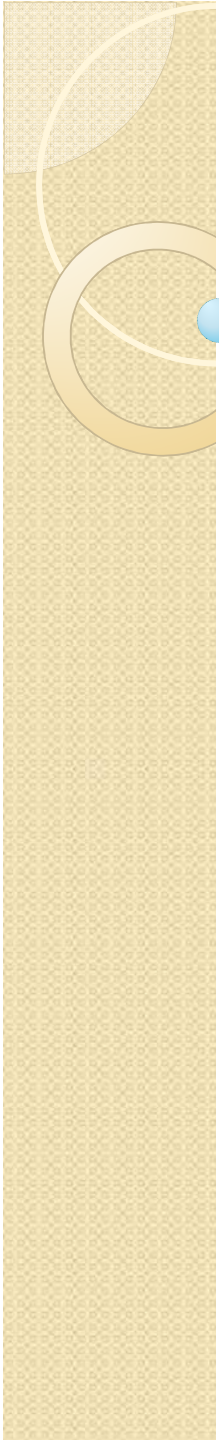
- Infrastruktūros plėtra

IV

- Bazinių kalbos išteklių ir įrankių kūrimas

V

- Sintaksinės ir semantinės analizės tobulinimas



Děkoju už děmesj