

DATORLINGVISTIKAS KURSI Latvijas Nacionālajā bibliotēkā

Datorlingvistika (*Computational Linguistics*) ir starpdisciplināra zinātņu nozare, saistīta ar dabiskās valodas datorapstrādes aspektiem. Datorlingvistikas uzmanības centrā ir dažādu lingvistisko fenomenu datorapstrādes modeļu izveide, kā arī dažādu praktisku dabiskās valodas apstrādes uzdevumu veikšana, izmantojot mūsdienīgus datortehnoloģiju rīkus. Datorlingvistika ietver sevī tādas lietojuma virzienus kā parsēšana (*parsing*), gramatikas indukcija (*grammar induction*), informācijas izguve un mašīntulkošana, atbalsts dažādiem semantiskā tīmekļa lietojumiem, kā arī tekstu korpusā balstītus pētījumus.

Ievadkursa saturs tiks balstīts gan uz datorlingvistikas formālajiem atzinumiem, gan arī uz tās praktisku lietojumu.

Kursu norise plānota 3 blokos:

I

- Ievads datorlingvistikā;
- Automātu teorija, automāti un gramatikas (*Automatons and grammars*);
- Kontekstneatkarīgās gramatikas un parsēšana (*Context free grammars and parsing*);
- Statistiskās metodes (*Statistical methods*);
- Marķēšanas metodes un principi (*Part-of-speech tagging*);
- Informācijas izguve (*Information retrieval*).

Kursi paredzēti apmeklētājiem bez vai ar nelielām priekšzināšanām.

Vieslektori:

Vībke Petersena (*Wiebke Petersen*) studējusi matemātiku un filozofiju Diseldorfas universitātē Vācijā. Ieguvusi maģistra grādu matemātikā Utrehtas universitātē Nīderlandē un doktora grādu datorlingvistikā – Diseldorfas universitātē. Disertācijas ietvaros veikusi Pānini klasiskā sanskrita gramatikas matemātisko analīzi. Šobrīd strādā pie projekta, kas saistīts ar rāmējumu matemātisko modelēšanu (*mathematical modelling of frames*). Viņas akadēmisko interešu lokā ir matemātiskā lingvistika un teorētiskā datorlingvistika.

Pāvels Sirotkins (*Pavel Sirotkin*) studējis vispārējo lingvistiku, kā arī angļu valodu un informācijas zinātņi Diseldorfas universitātē Vācijā. Maģistra grādu ieguvis vispārējā lingvistikā Oksfordas universitātē Lielbritānijā. Šobrīd strādā pie disertācijas Informācijas zinātnē Diseldorfas universitātē.

Viņa akadēmisko interešu laukā ir statistiskā dabiskās valodas apstrāde (*statistical natural language processing*) un tīmekļa meklēšanas novērtēšana (*web search evaluation*).

II

- Datorlingvistika, vārdu daudznozīmības problēma (*Computational linguistics / The role of WSD in CL*);
- Vārdu daudznozīmības problēmas risinājumi un daudznozīmība dabiskās valodas korpusā (*Word sense disambiguation / The role of corpora in WSD*);
- Vārdu daudznozīmības problēmas risinājumu iespējamā loma digitālo objektu pārvaldībā (*WSD's possible role in Digital Object Management*);
- Korpusa veidošanas praktiskie aspekti (*Corpus hands on - practical applications*);
- Vārdu daudznozīmības problēmas risinājumi digitālajās bibliotēkās un semantiskajā tīmeklī (*WSD issues re Digital Libraries / Semantic Web etc.*)

Vieslektors:

Stīvs Legrāns (*Steve Legrand*) ieguvis diplomu Vides zinātnē Brisbenas universitātē Austrālijā, diplomu praktiskajā programmēšanā (*Applied Computing*) Rokhemptonas universitātē Austrālijā, maģistra grādu digitālajos medijos, kā arī doktora grādu datorzinātnē un informācijas sistēmās Jiveskiles universitātē Somijā. Pēdējos 20 gadus darbojies visdažādākajās informācijas tehnoloģijas jomās un projektos, kas saistīti ar datorlingvistiku un programmēšanu Somijā, Meksikā, Lielbritānijā un Austrālijā. Viņa zinātnisko interešu loks saistīts ar daudznozīmības problēmas analīzi un risinājumiem, kā arī ar korpusa lingvistiku.

III

- Ievads korpusu lingvistikā (*Corpus Linguistics – An introduction*);
- Lietotāji orientētas pieejas valodai (*Usage-Based Approaches to Language*);
- Multifaktoru gramatika – viņpus kolokācijām (*Multifactorial Grammar – beyond the collocation*).

Vieslektors:

Dilans Glinns (*Dylan Glynn*) ieguvis diplomu semiotikā un teorētiskajā lingvistikā Sidnejas universitātē Austrālijā, maģistra grādu praktiskajā lingvistikā (*Applied Linguistics*) Jaundienvidvelsas universitātē Velsā. Darbojies Šarla V Institutā Parīzē, taču 2005. gadā pārcēlies uz Lēvenas universitāti Beļģijā, kur ieguvis doktora grādu. Šobrīd Dilana Glinna profesionālā darbība norit gan Lēvenas universitātes Lingvistikas departamentā, gan arī Lundas universitātes (Zviedrija)

Valodas un literatūras centrā. Interesējas par valodas multidimensionālo struktūru, sākot ar prosodiju un morfoloģiju līdz pat semantikai un kultūrai. Viņa specifisko interešu lokā ietilpst gramatisko konstrukciju semantika, sinonīmija un polisēmija, kā arī gramatikas, metaforas, metonīmijas un pragmatikas savstarpējā saistība.