

RTU studiju kurss "Datorizētā semantika un pragmatika"

33000 Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte

Vispārējā informācija

Kods	ETH727
Nosaukums	Datorizētā semantika un pragmatika
Studiju kursa statuss programmā	Brīvās izvēles
Atbildīgais mācībspēks	Tatjana Smirnova - Doktors, Asociētais profesors
Mācībspēks	Marina Platonova - Doktors, Profesors Zane Seņko - Docents (praktiskais) Tatjana Menise - Habilitētais doktors, Docents Tatjana Kelebeka - Doktors, Asociētais profesors Sintija Petroviča-Kļaviņa - Doktors, Lektors Oksana Ivanova - Docents (praktiskais) Alīna Vagele-Kricina - Lektors
Apjoms daļās un kredītpunktos	1 daļa, 2.0 kredītpunkti, 3.0 EKPS kredītpunkti
Studiju kursa īstenošanas valodas	EN
Anotācija	Studiju kurss galvenokārt ir vērst uz padziļinātu un augsti specializētu zināšanu līmeņa kompetenču un prasņu attīstīšanu studentiem, kuri apgūst gan humanitāro zinātņu, gan starpdisciplināro STEM+ balstītu, gan informācijas tehnoloģiju studiju programmas. Līdztekus vispusīgam pārskatam par fundamentāliem jautājumiem, kas saistīti ar semantisko un pragmatisko datu izgūšanu, vākšanu, organizēšanu un apstrādi (semantiskie un tematiskie lauki, nozīmes atveidojums, nozīmes paplašināšana, jēdzienu kartēšana un ontoloģiju veidošana, diskursa un patiesumvērtības analīze), studenti iepazīsies ar jaunākajiem sasniegumiem dabiskās valodas apstrādes un tekstrades jomā, kas viņiem palīdzēs izveidot visaptverošu teorētisko ietvaru dažādu ar dabiskās valodas apstrādi un tekstradi saistītu uzdevumu veikšanai. Noklausoties studiju kursā sagatavotās video lekcijas, patstāvīgi apgūstot saistīto literatūru un realizējot praktisko uzdevumu izpildi, studenti iemācīsies organizēt un attēlot nestrukturētus un daļēji strukturētus tekstuālos datus turpmākai izmantošanai dažādās tekstizraces, informācijas meklēšanas un zināšanu pārvaldības lietojumprogrammās. Studenti apgūs kompozicionālās un distributīvās semantikas pamatus, mācīsies analizēt tēmu struktūru un attīstīs savas prasmes semantisko modeļu, t.i., semantisko tīklu, taksonomiju, ontoloģiju un zināšanu grafu veidošanā un eksistējošo modeļu pielāgošanā un pielietošanā. Pēc studiju kursa apguves studenti spēs profesionāli iesaistīties dažādu valodu tehnoloģiju izstrādes komandu darbā, veikt pētījumus digitālās semantikas un pragmatikas jomā un līdzdarboties dažādu ar zināšanu pārvaldību saistītu risinājumu izstrādē, kā arī radīt risinājumus sarežģītām problēmām ar ierobežotu definīciju, ietverot arī jauna satura un informācijas modificēšanu, pilnveidošanu, uzlabošanu un integrēšanu esošajās digitālās semantikas un pragmatikas zināšanās, lai radītu jaunas un oriģinālas idejas. Studenti spēs piedalīties leksikogrāfisko resursu radīšanā un pielāgošanā dažādu nozaru un individuālu MVU un jaunuzņēmumu vajadzībām, izstrādājot jaunus produktus un pakalpojumus ar augstu pievienoto vērtību.
Mērķis un uzdevumi, izteikti kompetencēs un prasmēs	Studiju kursa mērķi: - nodrošināt studentiem visaptverošas zināšanas par semantiskajām attiecībām, kategoriālo gramatiku un nozīmes atveides modeļiem, lai viņi spētu veidot, uzturēt un pielāgot digitālos leksikogrāfiskos (piemēram, WordNet vārdnīcas, semantiskos tīklus, ontoloģijas u.c.) un teksta apstrādes resursus; - attīstīt studentu kompetenci izvēlēties, pielietot un pielāgot atbilstošas valodas apstrādes tehnoloģijas semantisko un pragmatisko datu izrāci, vākšanai, organizēšanai, apstrādei, glabāšanai un reprezentēšanai; - attīstīt studentu kompetenci ierobežotās definīcijas semantisko un pragmatisko datu, resursu un zināšanu līdzveidošanā un kopradē. Studiju kursa uzdevumi: - pilnveidot studentu prasmes digitālo semantisko, pragmatisko un sintaktisko modeļu izstrādē, pielāgošanā un piemērošanā, lai attēlotu un analizētu vārdu nozīmi; - attīstīt studentu prasmes izprast attiecīgo valodu tehnoloģiju darbības princips un izmantot šīs tehnoloģijas, lai apstrādātu nestrukturētus un daļēji strukturētus semantiskos datus dažādiem lietojumiem; - attīstīt studentu prasmes izmantot zināšanu grafus, vārdu nozīmju taksonomijas un distributīvās semantikas modeļus, lai organizētu un attēlotu semantiskos un pragmatiskos datus; - attīstīt studentu kompetenci tekstrades rīku izmantošanā dažādām pielietojuma jomām; - attīstīt studentu spēju sadarboties, lai piedāvātu un izstrādātu risinājumus dažāda sarežģītības līmeņa starpdisciplināriem semantikā un pragmatikā saņētiem valodu tehnoloģiju izaicinājumiem ar ierobežotu definīciju, izmantojot digitālos rīkus un iespējamus tehnoloģiskos risinājumus, kā arī piemērot un pielāgot digitālo vidi personīgajām vajadzībām.
Patstāvīgais darbs, tā organizācija un uzdevumi	Studenti patstāvīgi iepazīstas ar plašu saistītās literatūras klāstu un veic dažādu, ar teksta apstrādi saistītu praktisko uzdevumu izpildi, kas nodrošinās profesionālas zināšanas un kompetences valodu tehnoloģiju izvēlē un pielietošanā, lai līdzdarbotos risinājumu izstrādē, pielietotu un pielāgotu atbilstošu programmatūru un tehnoloģijas tekstuālo datu apkopošanā, organizēšanā, attēlošanā, apstrādē un pārvaldībā.

Literatūra	<p>Obligātā. / Obligatory : Kejriwal, Mayank. Knowledge graphs : fundamentals, techniques, and applications /Mayank Kejriwal, Craig A. Knoblock, and Pedro Szekely., xxvii, 530 lpp. : ilustrācijas ; 24 cm</p> <p>Stefanowitsch, A. Corpus-based Approaches to Metaphor and Metonymy. In Stefanowitsch, A. and Gries, S. Th. (eds) Corpus-based Approaches to Metaphor and Metonymy. Berlin, Germany: Mouton de Gruyter, 2006. https://www.researchgate.net/publication/292744446_Corpus-based_approaches_to_metaphor_and_metonymy</p> <p>Hasan, Md. Transformers in Natural Language Processing 2022. https://www.researchgate.net/publication/364842359_Transformers_in_Natural_Language_Processing</p> <p>Patil, N. Understanding Transformers: A Breakthrough in Natural Language Processing 2023 https://medium.com/@nihal.patil1122/understanding-transformers-a-breakthrough-in-natural-language-processing-8c1f5858ace4</p> <p>O'Grady, William. The Natural Syntax of Local Coreference Frontiers in Psychology. 12, 2021. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.660296/full</p> <p>Kroeger, P. Basic Concepts in Information Structure: Topic, Focus, and Contrast GIALens 2017 Volume 11, No. 1 https://www.diu.edu/documents/gialens/Vol11-1/Kroeger-Basic-Concepts-Information-Structure.pdf</p> <p>Papildu. / Additional : Jackson H., Ze Amvela E. Words, Meaning and Vocabulary: An Introduction to Modern English Lexicology UK: Cromwell Press, 2000 Liu, Z., Lin, Y., Sun, M. Compositional Semantics In: Representation Learning for Natural Language Processing. Springer, Singapore. 2020. https://doi.org/10.1007 Robert D. Van Valin, Jr. Randy J. LaPolla . Syntax: Structure, meaning and function. CUP, 2002 Dowty, D.R. Word Meaning and Montague Grammar The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ <i>Journal of Symbolic Logic</i> 48 (2):501-502</p>
Nepieciešamās priekšzināšanas	Pamatzināšanas Python programmēšanas valodā, DigComp 3.-4. Līmenī atbilstoši Eiropas Iedzīvotāju digitālās kompetences ietvaram, angļu valodas zināšanas vismaz B2 līmenī saskaņā ar EKPVA.

Studiju kursa saturs

Saturs	Pilna un nepilna laika klātienēs studijas		Nepilna laika neklātienēs studijas	
	Kontakt stundas	Patstāv. darbs	Kontakt stundas	Patstāv. darbs
Kompozicionālā semantika.	3	5	0	0
Kategoriālā gramatika.	3	4	0	0
Vārda nozīmes analīze dabiskās valodas apstrādē un tekstrādē.	3	5	0	0
Distributīvā semantika.	3	5	0	0
Kompozicionālie distributīvie modeļi.	3	5	0	0
Nozīmes atveide dabiskās valodas apstrādē.	3	5	0	0
Tēlainās valodas izpēte dabiskās valodas apstrādē un tekstrādē.	3	5	0	0
Informācijas un nozīmes ieguve digitālajā vidē.	3	5	0	0
Transformatora tipa modeļi un kontekstuālie vārdu atveidojumi.	3	5	0	0
Tēmas struktūra un fokusa struktūra tekstrādē.	3	4	0	0
Testi.	2	0	0	0
Kopā:	32	48	0	0

Sasniedzamie studiju rezultāti un to vērtēšana

Sasniedzamie studiju rezultāti	Rezultātu vērtēšanas metodes
Studenti spēj apkopot, uzturēt un pielietot digitālos leksikogrāfijas un teksta apstrādes resursus.	Metodes: praktiski uzdevumi, testi, eksāmens. Studenti demonstrē visaptverošas zināšanas par semantiskajām attiecībām, kategoriālo gramatiku un nozīmes atveides modeļiem, veiksmīgi izpildot ar dabiskās valodas apstrādi saistītus praktiskus uzdevumus.
Studenti spēj izvēlēties, pielietot un pielāgot atbilstošas valodu tehnoloģijas, kas izmantojamas dažādiem tekstizracēs, informācijas meklēšanas un zināšanu pārvaldības lietojumiem.	Metodes: praktiski uzdevumi, testi, eksāmens. Studenti demonstrē padziļinātu kompetenci darbā ar attiecīgo programmatūru un tehnoloģijām.
Studenti spēj apstrādāt, organizēt un attēlot nestrukturētus un daļēji strukturētus semantiskos datus, izstrādājot funkcionālus semantiskos un pragmatiskos modeļus.	Metodes: praktiski uzdevumi, testi, eksāmens. Studenti izstrādā funkcionālus semantiskos un pragmatiskos modeļus, taksonomijas, zināšanu grafus, pielieto WordNet vārdnīcu vārdu analīzei, kā arī veido semantiskos un sintaktiskos kokus.
Studenti spēj līdzveidot un līdzradīt, pielietot un pielāgot semantiskus un pragmatiskus datus, resursus un zināšanas specifiskā kontekstā.	Metodes: praktiski uzdevumi, testi, eksāmens. Studenti izvēlas piemērotus datus, resursus un rīkus un pielāgo tos savām pētniecības vai nozares vajadzībām.

Studenti spēj līdzdarboties starpdisciplināros projektos, kas saistīti ar dabiskās valodas apstrādes un tekstrades tehnoloģiju izstrādi, pielāgošanu un pielietošanu.	Metodes: praktiski uzdevumi, testi, eksāmens. Studenti var veiksmīgi līdzdarboties, piedāvājot un radot risinājumus starpdisciplināriem semantikā un pragmatikā sakņotiem valodu tehnoloģiju izaicinājumiem.
---	--

Studiju rezultātu vērtēšanas kritēriji

Kritērijs	% no kopējā vērtējuma
Praktiski uzdevumi	30
Testi	30
Eksāmens	40
Kopā:	100

Studiju kursa plānojums

Daļa	KP	Stundas			Pārbaudījumi			Brīvās izvēles pārbaudījumi		
		Lekcijas	Prakt d.	Laborat	Ieskaite	Eksām.	Darbs	Ieskaite	Eksām.	Darbs
1.	2.0	1.0	1.0	0.0		*				