

## RTU studiju kurss "Informācijas izguve"

33000 Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte

**Vispārējā informācija**

Kods	DE0939
Nosaukums	Informācijas izguve
Studiju kursa statuss programmā	Obligātais/Ierobežotās izvēles
Atbildīgais mācītbspēks	Marina Uhanova - Doktors, Asociētais profesors
Apjoms daļās un kredītpunktos	1 daļa, 6.0 kredītpunkti
Studiju kursa īstenošanas valodas	LV, EN
Anotācija	Studiju kursā tiek aplūkotas informācijas izguves sistēmu uzbūves principi, kā arī problēmas, kas saistītas ar lielu dokumentu kolekciju apstrādi. Studiju kursa ietvaros tiek izskatītas tēmas, kas saistītas ar tādiem informācijas izguves modeļiem kā vektoru telpas modelis, varbūtību modelis un valodas modelis, kā arī tiek aplūkoti kvantitatīvi izguves sistēmu novērtēšanas kritēriji. Tiek izskatītas meklēšanas veikspējas uzlabošanai paredzētas indeksu struktūras, kuras var izmantot ne tikai informācijas izguves sistēmās, bet arī datubāžu sistēmu izstrādē.
Mērķis un uzdevumi, izteikti kompetencēs un prasmes	Studiju kursa mērķis ir iepazīstināt studentus ar principiem un metodēm, kas tiek izmantotas informācijas izguves sistēmās. Studiju kursa uzdevumi ir: - nodrošināt studentiem zināšanas par informācijas izguves sistēmu veidošanas konceptiem un algoritmiem; - iemācīt studentus implementēt efektīvas indeksu struktūras, kas paredzētas meklēšanas veikspējas uzlabošanai gan strukturētās, gan nestructurētās datu kolekcijās; - nodrošināt studentiem zināšanas un prasmes lietot informācijas izguves sistēmu novērtēšanas kvantitatīvos kritērijus.
Patstāvīgais darbs, tā organizācija un uzdevumi	Patstāvīgais darbs sastāv no metodisko materiālu un literatūras apgūšanas, kā arī laboratorijas darbu izpildes. Katrā nodarbībā studenti saņem uzdevumu praktiskai realizācijai. Studenti uzsāk uzdevuma praktisko realizāciju tuvākajā laboratorijas darba laikā un, ja nepaspēj datorklasē, pabeidz to patstāvīgi.
Literatūra	Obligātā/Obligatory: 1. Cheng Xiang Zhai, Sean Massung. Text Data Management and Analysis. A Practical Introduction to Information Retrieval and Text Mining. ACM Books, 2016. 2. C. Manning, P. Raghavan, and H. Schütze. Introduction to Information Retrieval. Cambridge University Press, 2008. Papildu/Additional: 1. Liena Wiese. Advanced Data Management. Berlin; Boston: De Gruyter, Oldenbourg, 2015. 2. Kleppmann, Martin. Designing data-intensive applications. O'Reilly Media, 2017.
Nepieciešamās priekšzināšanas	Kopu un varbūtību teorija, programmēšanas valodas, datu struktūras.

**Studiju kursa saturs**

Saturs	Pilna un nepilna laika klātienes studijas		Nepilna laika neklātienes studijas	
	Kontakt stundas	Patstāv. darbs	Kontakt stundas	Patstāv. darbs
Strukturēti un nestructurēti dati. Lielo datu veidi un apstrādes principi.	4	2	0	0
Informācijas izguve, tās lietojumi un īpatnības.	4	6	0	0
Invertēta indeksa izveide un saspiešana.	6	10	0	0
Indeksu struktūras. B-koki, paplašināmās un lineāras jaucējtabulas.	8	12	0	0
Vārdnīcas un tolerantā informācijas izguve.	8	12	0	0
Vektoru telpas modelis un dokumentu ranžēšana.	6	10	0	0
Varbūtību modeļi un valodas modeļi informācijas izguvei.	8	12	0	0
Vaicājumu izpildīšana.	6	10	0	0
Informācijas izguves sistēmu novērtēšana.	6	10	0	0
Meklēšana tīmeklī. Saišu analīze. PageRank algoritms.	4	6	0	0
Ieteikumu sistēmas.	4	6	0	0
<b>Kopā:</b>	<b>64</b>	<b>96</b>	<b>0</b>	<b>0</b>

**Sasniedzamie studiju rezultāti un to vērtēšana**

Sasniedzamie studiju rezultāti	Rezultātu vērtēšanas metodes
Spēj izskaidrot strukturēto un nestructurēto datu glabāšanas un apstrādes principus.	Sekmīgi nokārtoti kontroldarbi un eksāmens. Kritēriji: students demonstrē izpratni par jomas tipisko problēmu būtību un to risināšanas metodēm.

Spēj pielietot konceptus un algoritmus, kas tiek lietoti informācijas izguves sistēmās.	Sekmīgi nokārtoti kontroldarbi un eksāmens. Sekmīgi izpildīti un aizstāvēti laboratorijas darbu uzdevumi. Kritēriji: spēj izstrādāt programmatūru uzdevuma risināšanai.
Spēj izstrādāt programmas, kas pielieto efektīvas indeksēšanas metodes strukturētu un nestrukturētu dokumentu kolekciju apstrādei.	Sekmīgi izpildīti un aizstāvēti laboratorijas darbu uzdevumi. Kritēriji: spēj izstrādāt algoritmu un programmatūru uzdevuma risināšanai.
Spēj pielietot kvantitatīvos kritērijus informācijas izguves sistēmu novērtēšanai.	Sekmīgi nokārtoti kontroldarbi un eksāmens. Kritēriji: students demonstrē kompetenci iegūto rezultātu interpretēšanā un tālākā apstrādē.

**Studiju rezultātu vērtēšanas kritēriji**

Kritērijs	% no kopējā vērtējuma
Laboratorijas darbu uzdevumi	50
Kontroldarbi	30
Eksāmens	20
Kopā:	100

**Studiju kursa plānojums**

Daļa	KP	Stundas			Pārbaudījumi		
		Lekcijas	Prakt d.	Laborat	Ieskaite	Eksām.	Darbs
1.	6.0	32.0	0.0	32.0		*	