

RTU studiju kurss "Datu integrācijas tehnoloģijas"

33000 Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte

Vispārējā informācija

Kods	DE0835
Nosaukums	Datu integrācijas tehnoloģijas
Studiju kursa statuss programmā	Obligātais/Ierobežotās izvēles
Atbildīgais mācītbspēks	Jānis Kampars - Doktors, Asociētais profesors
Apjoms daļās un kredītpunktos	1 daļa, 3.0 kredītpunkti
Studiju kursa īstenošanas valodas	LV, EN
Anotācija	Mūsdienās pastāvīgi tiek ģenerēti liela apjoma dati, kuru apstrāde ir teorētiski un tehnoloģiski sarežģīts process. No datiem potenciāli iegūstamais labums pieaug, ja dati no heterogēniem avotiem tiek integrēti un apstrādāti tuvu reālajam laikam, tādējādi minimizējot iegūtās informācijas un zināšanu latentumu. Šiem mērķiem izmanto datu straumēšanas, konteksta apstrādes un sistēmu adaptācijas tehnoloģijas. Studiju kursā tiek aplūkotas datu integrācijas tehnoloģijas, galveno vērību pievēršot datu straumju apstrādes un integrācijas tehnoloģijām, piemēram, Apache Spark un Apache Kafka. Tās tiek apskatīts datu dzīves cikla ietvaros, kas iekļauj datu integrāciju, apstrādi, interpretāciju un iegūtās informācijas izmantošanu sistēmu pielāgošanai tuvu reālajam laikam. Datu straumju integrācijā vienlīdz svarīga nozīme ir datu loģiskajai integrācijai un infrastruktūras risinājumiem, kam jānodrošina integrācijas loģikas izpilde dalītā, horizontāli mērogojamā vidē. Studiju kursa laikā tiek aplūkoti tuvu reālā laika straumju integrācijas lietošanas gadījumi, kas balstās uz Apache Spark, Apache Kafka, Apache Cassandra, Docker un CloudStack platformām.
Mērķis un uzdevumi, izteikti kompetencēs un prasmēs	Studiju kursa mērķis ir apgūt zināšanas un prasmes, kas nepieciešamas mērogojamu reāllaika datu integrācijas risinājumu izstrādei un darbināšanai mākoņdatošanas vidē.
Patstāvīgais darbs, tā organizācija un uzdevumi	Patstāvīgā darba uzdevums ir izveidot un eksperimentāli izvērtēt reāllaika datu integrācijas risinājumu uz esošas datu integrācijas platformas bāzes
Literatūra	Kampars, J., Grabis, J.: Near real-time big-data processing for data driven applications, In: Proceedings of the 3rd International Conference on Big Data Innovations and Applications, Innovate-Data 2017, Czech Republic, Prague, 21-23 August, 2017. Piscataway: IEEE, pp.35-42 (2017) Kleppmann, M. (2017) Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, O'Reilly Media Grabis, J., Stirna, J.: Advanced Context Processing for Business Process Execution Adjustment. In: Advanced Information Systems Engineering Workshops: CAiSE 2015 International Workshops, Sweden, Stockholm, 8-9 June, 2015. Springer International Publishing, pp.15-26 (2015)
Nepieciešamās priekšzināšanas	Datu bāzu vadības sistēmas, operētājsistēmas, datortīkli un datu analīze

Studiju kursa saturs

Saturs	Pilna un nepilna laika klātienes studijas		Nepilna laika neklātienes studijas	
	Kontakt stundas	Patstāv. darbs	Kontakt stundas	Patstāv. darbs
Datu dzīves cikls	4	0	0	0
Datu straumju integrācijas veidi: integrācijas pakešu un reāllaika režīmā	4	0	0	0
Datu straumēšana un reāllaika integrācijas risinājumi	4	0	0	0
Datu straumju integrācija loģiskā līmenī	4	0	0	0
Datu straumju integrācijas infrastruktūra	4	0	0	0
Datu straumju integrācijas mērogošanas risinājumi	4	0	0	0
Datu virzīta sistēmu adaptācija	4	0	0	0
Datu straumju integrācijas gadījumu izpēte	4	0	0	0
Kopā:	32	0	0	0

Sasniedzamie studiju rezultāti un to vērtēšana

Sasniedzamie studiju rezultāti	Rezultātu vērtēšanas metodes
Spēja izvēlēties piemērotu datu straumju integrācijas tehnoloģiju	Eksāmens
Spēja definēt datu integrācijas risinājumu loģiskajā līmenī	Eksāmens un patstāvīgais darbs
Spēja apvienot datu straumes	Patstāvīgais darbs

Studiju kursa plānojums

Daļa	KP	Stundas			Pārbaudījumi		
		Lekcijas	Prakt d.	Laborat	Ieskaite	Eksām.	Darbs
1.	3.0	16.0	0.0	16.0		*	