

RTU studiju kurss "Lielie dati un kultūras analītika"**33000 Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte****Vispārējā informācija**

Kods	DE1087
Nosaukums	Lielie dati un kultūras analītika
Studiju kursa statuss programmā	Obligātais/Ierobežotās izvēles; Brīvās izvēles
Atbildīgais mācītspēks	Tatjana Menise - Habilitētais doktors, Docents
Mācītspēks	Tatjana Smirnova - Doktors, Asociētais profesors
Apjoms daļās un kredītpunktos	1 daļa, 6.0 kredītpunkti
Studiju kursa īstenošanas valodas	LV, EN
Anotācija	Studiju kurss iepazīstina studējošos ar kultūras analītikas konceptuālajām, metodoloģiskajām un kritiskajām dimensijām. Kursā tiek aplūkots, kā lielo datu pārveido mūsu izpratni par tekstiem, naratīviem un kultūras modeļiem. Studējošie apgūs digitālās metodes humanitārajās zinātnēs, attīstot prasmes analizēt datus gan kvantitatīvi, gan kvalitatīvi, kā arī izpratni par ar datu pētniecību saistītajiem ētiskajiem un epistemoloģiskajiem jautājumiem. Studiju kursa ievadā tiek piedāvāts kritisks un konceptuāli plašs skatījums uz lielajiem datiem, pētot, kā datu infrastruktūras reprezentē un strukturē pasauli. Tiks analizēti dažādi tiešsaistē pieejami lielo datu resursi un atvērtās piekļuves korpusi, apspriežot to struktūru, pieejamību un ierobežojumus. Studiju kursa turpinājumā tiek apskatītas konkrētas analītiskās metodes, tostarp sentimenta analīze, tematu modelēšana, TF-IDF, mašīnmācīšanās pieejas un integrētas datu apstrādes metodes. Šīs metodes tiks pielietotas sarežģītu kultūras parādību izpētē, piemēram, naida runas, sazvērestības naratīvu, literāro žanru, digitālās folkloras, politiskā diskursa, kā arī mākslinieciskās radīšanas un uztveres modeļu analīzē. Studenti tiks iepazīstināti ar metodēm, izmantojot, izmantojot gadījumu izpēti, kas balstītas kultūras materiālos, un to pavadīs diskusijas par ētiskajiem riskiem, potenciālajiem aizspriedumiem un interpretācijas izaicinājumiem.
Mērķis un uzdevumi, izteikti kompetencēs un prasmēs	Studiju kursa mērķis ir iepazīstināt ar kultūras analītikas principiem un praksēm, uzsverot mijiedarbību starp kvantitatīvajām metodēm un kvalitatīvo interpretāciju. Studiju kursa uzdevumi ir attīstīt studējošo spējas: - atpazīt tādu datorizētu metožu kā sentimenta analīzes, tematiskās modelēšanas un mašīnmācīšanās potenciālu un ierobežojumus tekstu, naratīvu un citu kultūras datu formu pētniecībā; - kritiski lasīt un interpretēt sarežģītus pētījumus, kuros izmantotas šīs metodes, vērtējot to teorētiskos pieņēmumus, datu dizainu un analītisko pamatotību; - novērtēt un izstrādāt labi strukturētus korpusus, identificējot aizspriedumus, reprezentācijas trūkumus un ētiskos izaicinājumus datu iegūšanā un analīzē; - integrēt kvantitatīvos rezultātus ar kvalitatīvu interpretāciju, lai veidotu niansētas analīzes par kultūras parādībām, diskursiem un digitālās komunikācijas modeļiem.
Patstāvīgais darbs, tā organizācija un uzdevumi	Studiju kurss sāksies ar tiešsaistē pieejama korpusa kritisku izvērtējumu, patstāvīgi sagatavojot īsu rakstisku analīzi, kurā tiek vērtēts tā struktūra, reprezentativitāte, pieejamība un iespējamie aizspriedumi, kā arī pārdomāts, kā šādi dati var ietekmēt pētniecības rezultātus humanitārajās zinātnēs. Studējošiem būs jāizstrādā projekts, kurā viņi pielietos vienu vai vairākas kultūras analītikas metodes, piemēram, sentimenta analīzi, tematu modelēšanu vai mašīnmācīšanos izvēlētam datu kopumam un prezentēs savu projektu nodarbības laikā.
Literatūra	Obligātā. // Obligatory: Boyd, Danah, and Kate Crawford. "Critical Questions for Big Data." Information, Communication & Society 15, no. 5 (2012): 662–679. Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. Text as Data: A New Framework for Machine Learning and the Social Sciences. University of Chicago Press, 2019. Jockers, Matthew L. Macroanalysis: Digital Methods and Literary History. University of Illinois Press, 2013. Kitchin, Rob. The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences. Sage, 2014. Manovich, Lev. Cultural Analytics. MIT Press, 2020. Noble, Safiya Umoja. Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press, 2018. Olteanu, Alexandra, et al. "Social Data: Biases, Methodological Pitfalls, and Ethical Challenges." Frontiers in Big Data 3 (2020). Gitelman, Lisa, ed. Raw Data Is an Oxymoron. MIT Press, 2013. Papildu. // Additional: Glauner, Patrick, Petko Valtchev, and Sindi State. "Impact of Biases in Big Data." In Proceedings of the 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2018. Siddiqui, Nabeel. Cultural Analytics in R: A Tidy Approach. Springer, 2025. Hayles, N. Katherine, and Stefan P. Helmreich, eds. Switching Codes: Thinking Through Digital Technology in the Humanities and the Arts. University of Chicago Press, 2010.
Nepieciešamās priekšzināšanas	Angļu valodas prasmes vismaz B2 līmenī atbilstoši CEFR, ja studiju kurss tiek apgūts angļu valodā. Studējošiem jāspēj veikt arī pamatlīmeņa skaitļošanas eksperimentus, izmantojot tādas metodes kā TF-IDF, sentimenta analīze un tematu modelēšana.

Satus	Pilna un nepilna laika klātienes studijas		Nepilna laika neklātienes studijas	
	Kontakt stundas	Patstāv. darbs	Kontakt stundas	Patstāv. darbs
Lielie dati un paradigmas maiņa. Kā lielo datu revolūcija ir pārveidojusi zināšanas, kultūru un pētniecību dažādās jomās.	5	7	0	0
Kultūras analītikas ideja. Kultūras analītikas izcelsme un pieeja kultūras artefaktiem, praksēm un naratīviem kā datiem.	5	7	0	0
Lielo datu aizspriedumi un slēptie riski. Kā datu infrastruktūras un algoritmi reproducē sociālos, lingvistiskos un kultūras aizspriedumus, īpaši mazās un marginalizētās valodu kopienās.	5	7	0	0
Kvantitatīvās un kvalitatīvās metodes kultūras datos. Skaitliskās analīzes un interpretatīvo humanitāro pieeju savstarpējās saistības un papildināmības izpēte.	5	7	0	0
Mūsdienu pasaule kā korpusu multivisums. Datu kopu un arhīvu daudzveidības izpratne – no sociālajiem medijiem un digitālās komunikācijas līdz zinātnes, mākslas un ikdienas datiem – un to loma pētniecības perspektīvu un kultūras interpretāciju veidošanā.	5	7	0	0
Teksta reprezentācija datu veidā. Valodas pārvēršana analizējamos datos: vārdu frekvence, konkordance un teksta iezīmes.	5	7	0	0
Sentiments un emocijas kultūras tekstos. Afektu un polarizācijas mērīšana, izmantojot sentimenta analīzi un emociju leksikonu daudzvalodu un multikulturālās vidēs.	6	10	0	0
Tematu modelēšana un kultūras atklāsmes. Skaitļošanas modeļu izmantošana tēmu un diskursu noteikšanai lielapjoma kultūras datos.	6	10	0	0
Mašīnmācīšanās un kultūras procesi. Ievads klasifikācijā, klasterēšanā un prognozēšanā kā kultūras interpretācijas un raksturīguma analīzes rīkos.	6	10	0	0
Integrētas analītiskās datu apstrādes ķēdes. Dažādu skaitļošanas metožu plānošana un kombinēšana – no priekšapstrādes un iezīmju izgūšanai līdz vizualizācijai un interpretācijai – sarežģītu kultūras parādību izpēte.	6	10	0	0
Risinājumi mazajām valodām un minoritāšu diskursiem lielo datu laikmetā. Metodoloģisko, ētisko un infrastruktūras izaicinājumu izpēte, analizējot nepietiekami pārstāvētas valodas un diskursus lielapjoma digitālajās vidēs.	3	4	0	0
Melnās kastes problēma. Izaicinājumi izprast, kā modeļi apstrādā, nosaka prioritātes un rada rezultātus, kā arī šīs problēmas ietekme uz interpretāciju un atbildību kultūras pētniecībā.	3	5	0	0
Pārbaudes darbi.	4	7	0	0
Kopā:	64	98	0	0

Sasniedzamie studiju rezultāti un to vērtēšana

Sasniedzamie studiju rezultāti	Rezultātu vērtēšanas metodes
Spēj gūt praktisku pieredzi un demonstrēt izpratni par tādām metodēm kā sentimenta analīze, tematu modelēšana, mašīnmācīšanās un TF-IDF. Spēj lietot šos rīkus tekstu, naratīvu un kultūras parādību izpētē, kritiski izvērtējot to epistemoloģiskos pieņēmumus, interpretācijas ierobežojumus un ētiskās sekas humanitāro zinātņu kontekstā.	Diskusijas nodarbībās, korpusa kritisks vērtējums (rakstisks uzdevums), noslēguma projekts (prezentācija + rakstisks ziņojums), eksāmens. Studējošie piedalās diskusijās, demonstrējot spēju argumentēti skaidrot metodoloģiskās izvēles, kritiski analizēt piemērus un reflektēt par kultūras analītikas metodēm. Diskusiju sniegums tiks vērtēts, ņemot vērā ieguldījumu kvalitāti, spēju iesaistīties kritiski un konstruktīvi, kā arī pastāvīgu līdzdalību.
Spēj sasaistīt skaitļošanas analīzi ar humanitāro interpretāciju, sintezējot datu vadītu metožu rezultātus ar teorētisku, vēsturisku un kontekstuālu izpratni. Spēj interpretēt digitālos korpusus kā sarežģītus kultūras artefaktus, kurus veido lingvistiskie, sociālie un ideoloģiskie faktori.	Diskusijas nodarbībās, korpusa kritisks vērtējums (rakstisks uzdevums), noslēguma projekts (prezentācija + rakstisks ziņojums), eksāmens. Studējošie analizē tiešsaistē pieejamu korpusu, izvērtējot tā struktūru, reprezentativitāti, pieejamību, iespējamās aizspriedumus un nozīmi humanitārajā pētniecībā. Uzdevums tiks vērtēts, ņemot vērā analīzes precizitāti, kritiskās refleksijas dziļumu, kā arī rakstītā darba skaidrību un struktūru.
Spēj atpazīt, kā datu infrastruktūras, algoritmi un korpusi reproducē lingvistiskās, kultūras un reprezentācijas nevienlīdzības. Spēj izstrādāt un izvērtēt stratēģijas, lai veicinātu iekļaujošāku un atbildīgāku digitālo humanitāro zinātņu praksi, īpaši attiecībā uz mazajām valodām un minoritāšu diskursiem.	Diskusijas nodarbībās, korpusa kritisks vērtējums (rakstisks uzdevums), noslēguma projekts (prezentācija + rakstisks ziņojums), eksāmens. Studējošie pielieto vienu vai vairākas kultūras analītikas metodes, piemēram, sentimenta analīzi, tematu modelēšanu vai TF-IDF izvēlētam datu kopumam, sagatavo ziņojumu un prezentē rezultātus nodarbības laikā. Noslēguma projekts tiks vērtēts, ņemot vērā metožu pareizu pielietojumu, datu interpretācijas dziļumu un loģiskumu, kā arī prezentācijas skaidrību un profesionalitāti.

Studiju rezultātu vērtēšanas kritēriji

Kritērijs	% no kopējā vērtējuma
Korpusa kritisks vērtējums	30
Diskusijas nodarbībās	20
Noslēguma projekts un prezentācija	30

Eksāmens	20
Kopā:	100

Studiju kursa plānojums

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