

Databanken (C003803)

Wegens Covid19 kan mogelijk afgeweken worden van de onderwijs- en evaluatievormen. Dergelijke afwijkingen zullen via Ufora worden gecommuniceerd.

Cursusomvang *(nominale waarden; effectieve waarden kunnen verschillen per opleiding)*

Studiepunten 5.0 **Studietijd** 150 u **Contacturen** 45.0 u

Aanbodsessies en werkvormen in academiejaar 2020-2021

A (semester 2)	Engels	Gent	begeleide zelfstudie	12.5 u
			online hoorcollege	15.0 u
			online werkcollege: PC- klasoefeningen	17.5 u

Lesgevers in academiejaar 2020-2021

Bronselaer, Antoon	TW07	Verantwoordelijk lesgever
--------------------	------	---------------------------

Aangeboden in onderstaande opleidingen in 2020-2021

Master of Science in Statistical Data Analysis	stptn	aanbodssessie
	5	A

Onderwijstalen

Engels

Trefwoorden

Relationele databanken, SQL, Graafdatabanken, Data warehouses, Datakwaliteit

Situering

Het doel van deze cursus is het aanreiken van zowel theoretische kennis over als praktisch gebruik van databank technologie. De klemtoon ligt hierbij op het gebruik van bestaande databanken (interpreteren en lezen van schemas, data opvragen op een efficiënte manier, controle van de kwaliteit van data...).

Inhoud

- 1 Relationele databanken: het relationeel model, constraints, relationele algebra, de SQL bevragingstaal, gebruik van indices en query optimalisatie
- 2 NoSQL databanken: Document stores, Key-Value stores, graafdatabanken, "property graph" modellen, de Cypher bevragingstaal
- 3 Data warehousing: dimensionele modellen, ETL processen
- 4 Datakwaliteit: meten van datakwaliteit, edit regels in het Fellegi-Holt raamwerk, het foutlocatie probleem, minimale set covers, branch-and-bound oplossingen

Begincompetenties

Basis kennis van programmeren

Eindcompetenties

- 1 Begrijpen van het relationele model voor databanken en een relationele databank kunnen gebruiken
- 2 Begrijpen van NoSQL databanken in het algemeen en een graafdatabank kunnen gebruiken
- 3 Begrijpen van data warehouse modellen als een analytisch instrument
- 4 Begrijpen wat meten van datakwaliteit inhoudt en edit regels kunnen gebruiken in de praktijk

Creditcontractvoorwaarde

Toelating tot dit opleidingsonderdeel via creditcontract is mogelijk mits gunstige beoordeling van de competenties

Examencontractvoorwaarde

Dit opleidingsonderdeel kan niet via examencontract gevolgd worden

Didactische werkvormen

Begeleide zelfstudie, online hoorcollege, online werkcollege: PC-klasoefeningen

Toelichtingen bij de didactische werkvormen

De SQL oefeningen zullen worden aangeleerd via het Qexr oefeningenplatform dat toelaat om de SQL query taal grotendeels via zelfstudie aan te leren.

Belangrijk: omwille van COVID19 kunnen gewijzigde werkvormen uitgerold worden indien dit noodzakelijk blijkt.

Leermateriaal

- Slides
 - (Wetenschappelijke) artikelen
 - E-books
 - Korte filmpjes
- Geraamde totaalprijs: 20 euro

Referenties

- S. Abiteboul, R. Hull, V. Vianu, Foundations of databases, Addison Wesley, 1995
- T. De Waal, J. Pannekoek, S. Scholtus, Handbook of Statistical Data Editing and Imputation, Wiley, 2011
- R. Kimball, M. Ross, The Data Warehouse Toolkit (3rd edition), 201

Vakinhoudelijke studiebegeleiding

Evaluatiemomenten

periodegebonden en niet-periodegebonden evaluatie

Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode

Schriftelijk examen

Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode

Schriftelijk examen

Evaluatievormen bij niet-periodegebonden evaluatie

Verslag

Tweede examenkans in geval van niet-periodegebonden evaluatie

Examen in de tweede examenperiode is mogelijk

Eindscoreberekening

De niet-periodieke evaluatie zal 70% van de totaalscore bepalen, de periodieke evaluatie 30% van de totaalscore.

Om te slagen moet een student slagen voor beide evaluaties. Zoniet vervallen bovenvermelde gewichten. De totaalscore wordt dan berekend als het minimum van de scores van beide evaluaties.