UNIVERSITEIT GENT

Studiefiche

Academiejaar 2022-2023

Geavanceerde databanken (E018441)

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

Studiepunten 6.0	**Studietijd 180 u**	**Contacturen** 60.0u

Aanbodsessies en werkvormen in academiejaar 2022-2023

A (semester 1)	Nederlands	Gent	hoorcollege	30.0u
			werkcollege	30.0u

Lesgevers in academiejaar 2022-2023

Bronselaer, Antoon	TW07	Verantwoordelijk lesgever
De Tré, Guy	TW07	Medelesgever

Aangeboden in onderstaande opleidingen in 2022-2023	**stptn**	**aanbodsessie**
Master of Science in Business Engineering(afstudeerrichting Data Analytics)	6	A
Master of Science in Business Engineering(afstudeerrichting Operations Management)	6	A
Master of Science in Computer Science Engineering	6	A
Master of Science in de industriële wetenschappen: informatica	6	A
Master of Science in de informatica	6	A
Master of Science in de ingenieurswetenschappen: computerwetenschappen	6	A

Onderwijstalen

Nederlands

Trefwoorden

datakwaliteit, data warehousing, NoSQL, imputatie, normalisatie, indexering

Situering

Deze cursus is een specialisatiecursus waarin theorie en toepassingen rond databanken aan bod komen. Een eerste deel van de cursus behandelt methoden voor het meten en verbeteren van data kwaliteit. In dit deel komen verschillende algoritmen aan bod die toelaten om fouten te detecteren, alsook methoden om deze fouten systematisch te gaan verbeteren. Ook de verschillende normaalvormen van databanken en hun praktisch nut worden aangeleerd. In een tweede deel worden niet-relationele databankmodellen besproken (NoSQL) en leren we over het concept van data warehousing. Tot slot wordt ook aandacht besteed aan elementaire opslag van data en het sneller bevragen van data (indexering).

Inhoud

- Datakwaliteit: elementaire meettechnieken
- Edit regels
- Functionele afhankelijkheden en normaalvormen
- Herstel van systematische en willekeurige fouten
- Dynamische analyse van datakwaliteit
- Ontdubbeling van data
- Controlegetallen
- Data warehousing
- Primaire en secundaire bestandsorganisatie
- NoSQL-oplossingen en gedistribueerde databanken voor het beheer van Big Data

Begincompetenties

basisprincipes van (relationele) databanken

Eindcompetenties

1 Begrijpen van basistechnieken voor meten van datakwaliteit.
2 Begrijpen van basistechnieken voor verbetering van datakwaliteit.

3 De basisprincipes van fysieke databankopslag en indexering kennen.
4 'Datawarehouse'-technologie kunnen toepassen.
5 Weten hoe 'Big' data kunnen worden beheerd via NoSQL en gedistribueerde databanktechnologie.

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

Werkcollege, Hoorcollege

Toelichtingen bij de didactische werkvormen

Omwille van COVID19 kunnen gewijzigde werkvormen uitgerold worden indien dit noodzakelijk blijkt.

Leermateriaal

- Syllabus (Engelstalig). Bijkomend leermateriaal beschikbaar via de elektronische leeromgeving Ufora
- richtprijs: 20 EURO

Referenties

- R. Elmasri, S.B. Navathe, Fundamentals of Database Systems, Seventh Edition, Pearson Addison-Wesley, Boston USA, 2016 (ISBN: 9780133971330)
- T. de Waal, J. Pannekoek, S. Scholtus, Handbook of Statistical Data Editing and Imputation, Wiley, 2011 (ISBN: 978-0-470-54280-4)
- G. De Tré, Principes van databases, 2e editie. Pearson Education Benelux, 2017 (ISBN: 978-90-430-3580-4)

Vakinhoudelijke studiebegeleiding

Alle oefeningenlessen en het practicum worden begeleid door assistenten.

Evaluatiemomenten

periodegebonden en niet-periodegebonden evaluatie

Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode

Schriftelijk examen, Openboekexamen

Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode

Schriftelijk examen, Openboekexamen

Evaluatievormen bij niet-periodegebonden evaluatie

Verslag

Tweede examenkans in geval van niet-periodegebonden evaluatie

Examen in de tweede examenperiode is enkel mogelijk in gewijzigde vorm

Toelichtingen bij de evaluatievormen

Periodegebonden evaluatie: Schriftelijk examen met open boek. Indien het niet mogelijk is een "on campus" examen te organiseren, zullen de examens via MS Teams worden afgenomen.
Niet-periodegebonden evaluatie: ontwerpoefening met verslag/ontwerp in te dienen halverwege het semester.

Eindscoreberekening

Er wordt een gewogen gemiddelde van beide evaluaties met gewichten:
- Periodegebonden evaluatie (15/20)
- Niet-Periodegebonden evaluatie (5/20)

Studenten kunnen voor dit opleidingsonderdeel enkel slagen wanneer ze voor **beide** onderdelen minimum 10/20 behalen. Als studenten voor minstens één van de onderdelen minder dan 10/20 behalen dan gelden onderstaande regels:

- Indien men voor minstens één onderdeel van de evaluatie een 8/20 of 9/20 behaalt, kan men niet slagen voor het geheel van het opleidingsonderdeel. Als de eindscore via gewogen gemiddelde een cijfer van 10 of meer op 20 zou zijn, wordt dit teruggebracht tot het hoogste niet-geslaagd cijfer, nl. 9/20.
- Indien men voor minstens één onderdeel van de evaluatie minder dan 8/20 behaalt, kan men niet slagen voor het geheel van het opleidingsonderdeel. Als de eindscore via gewogen gemiddelde een cijfer van 8 of meer op 20 zou zijn, wordt dit teruggebracht tot het hoogste

niet-delibereerbare cijfer, nl. 7/20.

Een partiële vrijstelling voor de niet-periodegebonden evaluatie is mogelijk.

Faciliteiten voor werkstudenten

- De niet-periodieke evaluatie kan thuis worden gedaan en is een individuele activiteit. Studenten krijgen een server tot hun beschikking.
- Herhalingsoefeningen en oefeningen rond SQL functionaliteit nodig tijdens de niet-periodieke evaluatie kan worden ingeoefend via het Qexr platform.
- Korte kennisclips die de basis aspecten van de cursus samenvatten worden ter beschikking gesteld via Ufora.