GHENT UNIVERSITY

Course Specifications

Valid as from the academic year 2026-2027

Big Data Technology (E018241)

Course size *(nominal values; actual values may depend on programme)*

Credits 3.0 **Study time 90 h**

Course offerings in academic year 2026-2027

A (semester 1)	English	Gent

Lecturers in academic year 2026-2027

Colpaert, Pieter	TWO6	lecturer-in-charge
Gonçalves Crisóstomo Esteves, Beatriz	TWO6	co-lecturer
Rojas Melendez, Julian Andres	TWO6	co-lecturer

Offered in the following programmes in 2026-2027	crdts	offering
Master of Science in Bioinformatics(main subject Engineering)	3	A
Master of Science in Operations Research Engineering(main subject Manufacturing and Supply Chain Engineering)	3	A
Master of Science in Operations Research Engineering(main subject Transport and Mobility Engineering)	3	A
European Master of Science in Nuclear Fusion and Engineering Physics	3	A
Master of Science in Computer Science Engineering	3	A
Master of Science in Computer Science Engineering	3	A
Master of Science in Operations Research Engineering	3	A

Teaching languages

English

Keywords

Big Data platforms, data architecture, data governance, interoperability, trust, velocity, compliance.

Position of the course

This course addresses the technologies and design principles required to manage data at scale. It focuses on the challenges that emerge when data need to be distributed, heterogeneous, continuously generated, or when the data are subject to interoperability, trust, and compliance requirements. By covering data exchange, knowledge representation, stream-oriented architectures, and workflow orchestration, the course equips students with the conceptual foundation needed to understand and design scalable and responsible big data infrastructures.

Contents

- An opening guest lecture positioning big data in societal, industrial and/or strategic context
- Knowledge representation
  - Interoperability
  - Storage models
  - Serializations and data formats
  - Web APIs for data storage and query answering
- Foundations of Big Data
  - Volume, velocity, variety, veracity
  - Historic foundations and architectural concepts
  - Divide and conquer strategies for dealing with volume
- Veracity and trust
  - Verifiable Credentials
  - Identifiers and identity

    - Data spaces and contract negotiation
    - From access control to usage control with ODRL and DPV
  - Velocity and Streaming data
    - Stream processing
    - Publish-subscribe protocols
    - Architectures such as extract-transform-load, event sourcing, medaillon, shift-left, etc.
  - Compliance and governance
    - Intellectual property rights on data
    - European frameworks relevant to data
  - Variety and AI: dealing with unstructured data
  - Scheduling, orchestration, and workflows

  We include industry examples and lab sessions.

Initial competences

- basic programming skills: experience with Object Oriented Programming (passed the course Computer Programming E017210 or an equivalent course)
- elementary understanding about basic data formats (CSV, JSON, etc.)
- using the command-line

Final competences

1 Explaining the main challenges of big data systems in terms of among others volume, velocity, variety, and veracity.
2 Explaining distributed storage and processing architectures and the trade-offs.
3 Applying appropriate knowledge representation, storage, and serialization technologies for specific data-intensive use cases.
4 Explaining how interoperability can be achieved across heterogeneous data systems.
5 Explaining and applying trust mechanisms in data ecosystems, including identifiers, digital identity, verifiable credentials, and usage control.
6 Explaining and applying the legal and governance constraints that shape big data infrastructures, including privacy, intellectual property, and European data regulation.
7 Dealing with high-velocity data through stream processing, event-driven architectures, and messaging patterns.
8 Explaining how heterogeneous and unstructured data can be exploited in AI-oriented systems.
9 Designing and orchestrating workflows for data-intensive applications using workflow and integration technologies.
10 Assessing big data architectures with respect to scalability, interoperability, trustworthiness, and compliance.
11 Applying core big data concepts and technologies in practical settings.

Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

Conditions for exam contract

This course unit cannot be taken via an exam contract

Teaching methods

Lecture, Practical

Study material

None

References

- European Union Agency for Fundamental Rights and Council of Europe. Handbook on European data protection law. Publications Office of the European Union, 2018. ISBN 978-92-871-9849-5 978-92-9491-903-8 978-92-9491-901-4. doi: 10.2811/58814.
- Mbata, Anthony, Yaji Sripada, and Mingjun Zhong. "A survey of pipeline tools for data engineering." *arXiv preprint arXiv:2406.08335* (2024).

Course content-related study coaching

Assessment moments

end-of-term and continuous assessment

Examination methods in case of periodic assessment during the first examination period

Oral assessment open-book, Written assessment open-book

Examination methods in case of periodic assessment during the second examination period

Oral assessment open-book, Written assessment open-book

Examination methods in case of permanent assessment

Assignment

Possibilities of retake in case of permanent assessment

examination during the second examination period is possible in modified form

Extra information on the examination methods

- Periodical evaluation: multiple questions are solved in writing and submitted (open-book). The questions are also discussed orally.
- Non-periodical evaluation: graded lab session reports, possibly in groups.

Calculation of the examination mark

The final grade is the sum of the two parts: 12/20 for the exam (60%) and 8/20 for the non-periodical evaluation (40%).
In case the grade for any part is less than 10/20, the final grade is capped at 9/20.
In case the grade for any part is 7/20 or less, the final grade is capped at 7/20.
For students who have not passed the non-periodical evaluation, an alternative assignment is provided in the second examination period.