

Biological Databases (C004611)

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

Credits 3.0

Study time 90 h

Course offerings in academic year 2027-2028

A (semester 2)

English

Gent

Lecturers in academic year 2027-2028

Van Criekeinghe, Wim

LA26

lecturer-in-charge

Menschaert, Gerben

LA26

co-lecturer

Offered in the following programmes in 2027-2028

Master of Science in Bioinformatics(main subject Bioscience Engineering)

crdts

offering

3

A

Master of Science in Bioinformatics(main subject Systems Biology)

3

A

Master of Science in Biochemistry and Biotechnology

3

A

Exchange programme in Biochemistry and Biotechnology (master's level)

3

A

Exchange Programme in Bioinformatics (master's level)

3

A

Teaching languages

English

Keywords

Biological databases, BioSQL, XML, Linux, PHP, networking, web-based application, heterogeneous data analysis and integration

Position of the course

The enhanced course on bioinformatics focuses on database design and implementation of biological datasets. Both string based data (sequence) and integer/real data as obtained by genome-wide experiments are integrated in different datamodels. It complements the (bio)informatics background by applying modern application development environments to (relational) database methodologies.

Contents

Remark: Part 3 (Computer infrastructure) is omitted for course offering B

1. Biological database management
 - 1.1. The relational datamodel
 - 1.2. Data normalisation
 - 1.3. Structured query language
 - 1.4. BioSQL/Chado: a biological datamodel
 - 1.5. Object oriented databases
 - 1.6. Biological databases
 - 1.7. Biological database integration
 - 1.8. Distributed annotation systems (DAS)
 - 1.9. Hierarchical and frame-based system (XML, DAML+OIL)
2. Heterogeneous database integration
 - 2.1. (Application) integration frameworks
 - 2.2. Analysis methods: coupling large databases to statistics
 - 2.3. Non relational data integration using LWP, Bots & Spiders
 - 2.4. Text mining
 - 2.5. Alternative query systems
3. Computer infrastructure *(only for course offering A)*
 - 3.1. Basic computer architecture

- 3.2. Linux operating system (basic and advanced)
- 3.3. Packages and BioLinux
- 3.4. Networks (local, wan, protocols)
- 3.5. Hardware integration
- 3.6. High Performance Computer infrastructures
- 4. Ethical and sociological considerations regarding data storage and availability (to both the scientific community and patients themselves)

Initial competences

Basic knowledge of mathematics, computer science and bioinformatics.

Final competences

- 1 Being able to setup a propriatary database system, combining heterogeneous datasources and analyse/report in a web environment.
- 2 Master the use of AI to dessign and implement biological databases-- Klik om te editeren ---
- 3 Best practices in designing and maintaining SQL and noSQL data repositories. Understand modern data infrastructure including but not limited to semantic web technologies, blockchain, SPARQL, solid, ...- Klik om te editeren ---

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

Seminar, Lecture, Independent work

Extra information on the teaching methods

Theory: oral lectures

Exercises: link to online tutorials + exercises

Study material

Type: Slides

Name: Course slides + online material + tutorials

Indicative price: Free or paid by faculty

Optional: no

References

UFORA and GITHUB

Course content-related study coaching

UFORA

Assessment moments

end-of-term assessment

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

Assignment

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

Assignment

Examination methods in case of permanent assessment

Possibilities of retake in case of permanent assessment

not applicable

Extra information on the examination methods

Assignment with oral presentation followed by discussion

Students who eschew period aligned and/or non-period aligned evaluations for this course unit may be failed by the examiner.

Calculation of the examination mark

100% assignment (by 2 to 3 independent evaluators), following aspects will be evaluated:

- Topic + implementation

- Relational DB usage
- Reporting on biological data in light of research question
- Presentation