Curriculum

Biomedical data has exponentially grown in the last decade due to the advances in technologies for data generation and ingestion, e.g., liquid biopsies, medical imaging, and sensor-based wearable devices. Biomedical data can be represented using various data formats and access can be controlled following distinct regulations. This rich spectrum of biomedical data encodes knowledge about disease progression, drug mechanisms of actions, and genetic disorders that may contribute to a more accurate diagnosis and thoughtful treatment prescriptions. Using this wealth of knowledge demands skilled professionals for dealing with data coming from a range of biological and medical research and implementing innovative techniques for accurate knowledge discovery.

The BIOMEDAS curriculum builds upon the fields of:

- Computer Science: discipline of formalisms and scalable algorithmic processes;
- Data Science: pseudo-field for discovering intrinsic data properties, value, and actionable insights; and
- Biomedicine: area that combines natural sciences, especially the biological and physiological sciences, with clinical medicine

and thus, offers a multidisciplinary curriculum to train scientists with the required skills to address the challenges of transforming biomedical data into actionable knowledge which in turn will support discoveries and advance biomedicine.

Structure

The structure of the program is based on the HBRS regulations and the corresponding PhD regulations and is schematically shown in Figure 1.

The individual timetable is compiled by the thesis committee together with the doctoral candidate according to the prior knowledge of the respective candidate. The intention here is to align the accompanying program closely to the individual needs of the respective candidate with the aim of efficient further qualification.

To satisfy the requirements, the doctoral candidate has to submit a thesis and prove successful participation in a structured program of a minimum of 300 hours of lessons (1 hour of lesson corresponds to 45 minutes of time) as shown in Table 1. Participation in courses with relevant content from other graduate programs or university lectures of partner institutions may also be accepted upon request.

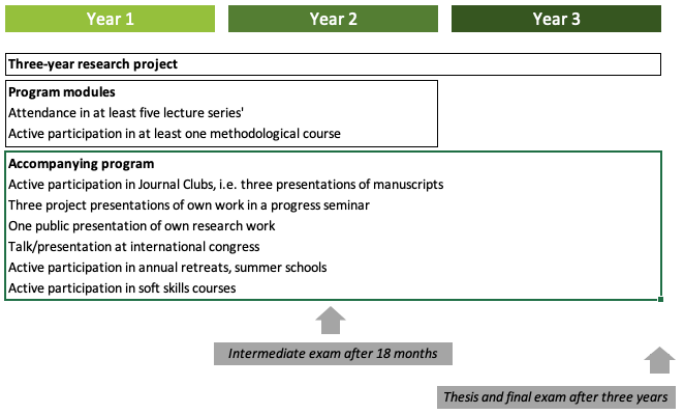

Figure Structure of the BIOMEDAS program.

Table 1 Exemplary structure of a doctoral candidate's individual timetable within BIOMEDAS

General specifications	Student's schedule (example)
Three-year research project	Research project selected
Program modules • Successful participation in **lecture series** [attendance in at least five lecture series'] • Active participation in **methodological courses** attendance in [at least one course]	• Attendance of six lecture series', each 7x1.5h → 84 lessons • Attendance of two courses, each 2 full days → 40 lessons
Accompanying program • Participation in a **Journal Club** with at least three presentations of publications • Participation in a **Progress Seminar** with presentations of own work • **Conference participation** with own contribution (oral or poster presentation) [at least one conference participation] • Active participation in **annual retreats** [attendance at at least one retreat] • Active participation in **soft skills courses** [attendance in at least one course]	• Attendance and contribution to JC, annually 20x45min, 3 years → 60 lessons • Attendance and contribution to PS, annually 20x45min, 3 years → 60 lessons • Attendance and contribution to an international congress, 2 full days → 20 lessons • Attendance and contribution to annual retreat, 2 times, full day → 20 lessons • Attendance of soft skills courses, 2 times, each full day → 20 lessons

Based on the provided curriculum, BIOMEDAS students will develop the following competencies, required to model biomedical problems and analyzing biomedical data using mathematical and computational approaches:

- Computational thinking by training computer science skills to address problem-solving effectively.
- Mathematical and statistical modelling and analysis to provide a universal language essential for producing reliable and reproducible results from real-world problems and data.
- Conceptual knowledge in translational bioinformatics, clinical informatics, consumer health informatics and public health informatics.
- Improving the ability to make ethical choices necessary for responsible conduct of research.

Program modules

The offered lectures and courses are organized in different clusters, which illuminate different areas and aim at diverse knowledge and skills.

Biological and Medical Concepts

The cluster “Computational Biology” is organized by Anke Kraft and Annett Ziegler.

The modules for this cluster aim at understanding the principles of biological and medical concepts incl. sample preparation, experimental design and data generation. The skills developed include:

- Introduction into basic concepts of biology, virology, microbiology, immunology and genetics
- Conceptual understanding of how samples are generated, experiments are designed and data generated to be able to assess the impact for data analysis

Computational Biology (CB)

The cluster “Computational Biology” is organized by Andreas Klötgen and David DeLuca.

The modules for this cluster aim at understanding of data format, representation and analysis in biomedical sciences and complex concepts for understanding the biological system as a whole. The skills developed include:

- Introduction into basic bioinformatics concepts & statistics
- Bioinformatics methods for data representation and visualization using networks and graphs
- Training in concepts and techniques for evolutionary analyses
- Assessing quality of bioinformatics data for reliable and robust data analysis
- Computational approaches for sequencing data analysis and integration, from alignment over filtering to downstream analyses
- Advanced approaches for complex data integration of various sequencing types and computationally predicted sequence features
- Conceptual understanding and practical introduction to systems biology approaches

Computational Method Development (CMD)

The cluster “Computational Method Development” is organized by Maria-Esther Vidal and Sören Auer.

The modules for this cluster aim at enhancing the competencies of computational thinking and mathematical and statistical modelling. The skills developed include:

- Mathematical languages to formalize biomedical problems and solutions.
- Algorithmic techniques to effectively solve biomedical problems.
- Computational data structures to efficiently store biomedical data.
- Data models and ontologies to formally represent biomedical knowledge.
- Data management techniques to devise pipelines to collect and integrate biomedical data.
- Visual literacy to devise meaningful and self-interpretable visual representations of the main features of biomedical data and knowledge.
- Data management techniques for assessing, tracking, and improving data quality.

Machine Learning and Data Mining (MLDM)

The cluster “Machine Learning and Data Mining” is organized by Frank Klawonn and Tim Kacprowski.

The primary objective of these modules is to enhance computational thinking and mathematical and statistical modelling of the students. The skills to be reached in these courses are:

- Statistical methods to analyse, describe, and curate biomedical data.
- Data analysis for feature selection in biomedical data.
- Artificial intelligence approaches to empower statistical techniques with the ability to learn and predict patterns from biomedical data.
- Machine learning methods for modelling predictive biomedical problems, as well as explaining the outcomes of these methods in biomedical problems.

Interdisciplinary (I)

The cluster “Interdisciplinary” is organized by Maria-Esther Vidal and Roland Seifert.

These modules aim to add overarching topics to the curriculum. The skills to be reached in these courses are:

- Principles that support ethical choices (Responsible conduct of research)
- Principles that guide the publication and management of research data
- Methods to conduct and report research studies to ensure reproducibility
- Open science and the development of services for research data management
- Formalisms based on Open Research Knowledge Graphs to transform document-based scientific literature review into knowledge-based information flows

BIOMEDAS Lectures and Courses

The following lectures and courses are specifically designed for the BIOMEDAS students.

Biological and Medical Concepts: Principles and Data Generation

No	Type	Title	Sem.	Organized by	Planned duration
BMC-L1	Lecture	**Mandatory**: Introduction to biological and medical concepts	3	Kraft and Ziegler	x x 1.5h

Computational Biology

No	Type	Title	Sem.	Organized by	Planned duration
CB-L1	Lecture	Introduction to bioinformatics data types and analysis techniques	1	McHardy / Klötgen and DeLuca with help of Jung, Xiao	14 x 1.5h
CB-C1	Course	Introduction to bioinformatics data types and analysis techniques	1	McHardy / Klötgen and DeLuca with help of Jung, Xiao	2 x 6h
CB-L2	Lecture	Comparative genomics, phylogenetics and other sequence-based analyses	3	Lauber	4-7 x 1.5h
CB-C2	Course	Comparative genomics, phylogenetics and other sequence-based analyses	3	Lauber	3-4 x 6h
CB-L3	Lecture	Advanced techniques for multi-omics data analysis and integration (oder?)	4	Li, Kacprowski and Xu	10 x 1.5h
CB-C3	Course	Advanced techniques for multi-omics data analysis and integration	4	Li, Kacprowski and Xu	2 x 6h

Computational Method Development

No	Type	Title	Sem.	Organized by	Planned duration
CMD-L1	Lecture	Introduction to Scientific Databases	1	Vidal	6 x 1,5h
CMD-C1	Course	Scientific database programming	4	Vidal	6 x 1.5h
CMD-L2	Lecture	Reproducibility (code versioning, documentation, data archiving)	2	Davenport	2 x 1,5h
CMD-L3	Lecture	Biomedical Data Management	3	Vidal	7 x 1.5h
CMD-L4	Lecture	Data Visualization	4	Klawonn	6 x 1.5h
CMD-L5	Lecture	Data structures and algorithm design	2	tba	tbd (6 x 1,5h)

Machine Learning and Data Mining

No	Type	Title	Sem.	Organized by	Planned duration
MLDM-L1	Lecture	Introduction to machine learning and data mining	2	Großhennig, Kacprowski, Jung, Deserno	14 x 1,5h
MLDM-L2	Lecture	An introduction to mechanistic mathematical modeling approach in biology and medicine	3	Meyer-Hermann	8 x 1.5h
MLDM-L3	Lecture or Seminar	Explainable Artificial Intelligence	4	tba (external Lecturer: Andreas Holzinger)	tbd (probably 2 x 6h)

Interdisciplinary

No	Type	Title	Sem.	Organized by	Planned duration
I-L1	Lecture	**Mandatory**: Ethics and Scientific Integrity	1	Seifert	3 x 1.5h
I-L2	Lecture	**Mandatory**: Responsible Data management	4	Vidal	7 x 1.5h
I-C1	Course	Research career development and Proposal writing	3	Auer	2 x 1.5h
I-C2	Course	Better scientific presentations and writing	2	Kacprowski / Deserno	6 x 1.5h

Lectures and courses offered in other programs

The following lectures and courses are offered via other programs, but are open to BIOMEDAS students upon request. As details such as schedule might change please always refer to the BIOMEDAS office.

Please note that this list is not complete and other courses may be taken as well, see also chapter 3.4.

Computational Biology

No.	Type	Title	Sem.	Organized by	Planned duration
CB-EL1	Lecture	Introduction to Molecular Data Science	1-2	DeLuca, Chouvarine	3 x 6h
CB-EC1	Course	Molecular Data Science with R	1-2	DeLuca, Chouvarine	3 x 6h
CB-EL2	Lecture	Systems Immunometabolism	1-6	Hiller	14 x 2h
CB-EC2	Course	Systems Immunometabolism	1-6	Hiller	2-3 weeks
CB-EL3	Lecture / Tutorials	Applied Bioinformatics	3-4	Hiller	14 x 1.5h
CB-EC3	Course	Applied Bioinformatics	3-4	Wegner	2 weeks
CB-EL4	Lecture / Tutorials	Systems biology	3-4	Wegner	14 x 1.5h
CB-EC4	Course	Systems biology	3-4	Wegner	2 weeks
CB-EL5	Lecture	Network Biology and Graph Theory	2	Kacprowski	TBD

Computational Method Development

No.	Type	Title	Sem.	Organized by	Planned duration
CMD-EL2	Lecture	Knowledge Engineering & Semantic Technologies	3	Auer	15 x 1.5h
CMD-EL3	Lecture	Big data computing, distributed computing and Hadoop	3	Höppner	tbd (12 x 3h)
CMD-EC1	Course	Programming Course Python	1-2	Hiller	1 week
CMD-EC2	Course	Knowledge Modeling	3	Auer	15 x 45min

Machine Learning and Data Mining

No.	Type	Title	Sem.	Organized by	Planned duration
MLDM-EL1	Lecture / Course	Machine learning and biomarker signatures	1-2	Hiller, O'Connor	14 x 2h

Interdisciplinary

No.	Type	Title	Sem.	Organized by	Planned duration
I-EL1	Lecture	**Mandatory** Good scientific practice	1	HBRS	3 x 1.5h

Accompanying program

Journal Club and Progress Seminar

A BIOMEDAS Journal Club as well as Progress Seminar will be established. BIOMEDAS students will present recent publications from the field (Journal Club) or their own work (progress seminar). An organizing committee (BIOMEDAS students) will be in charge of the schedule etc. Membership in the organizing committee will move every semester. By agreeing timely on the future committee, a smooth hand-over is ensured.

The Journal Club and Progress Seminar are hosted by the supervisor of the respective presenting PhD student.

Title	Lecturer/Organizer	Duration/Credit	Time/Place
BIOMEDAS Journal Club	BIOMEDAS students	45 min/bi-weekly	tba/web-based
BIOMEDAS Progress Seminar	BIOMEDAS students	45 min/bi-weekly	tba/web-based

Annual Retreat

A PhD retreat of the Program BIOMEDAS is taking place annually.

Soft Skill Courses

Please refer to the courses offered via the HBRS.

Alternative Courses

Students enrolled in BIOMEDAS are encouraged to attend courses with relevant content from other graduate programs or university lectures of partner institutions. Hours of lessons can be accredited after consulting with the thesis committee and in line with the requirements of the program.