# ***Syllabus:***

# ***Topics in Computational Biology & Analysis of Scientific Data using R***

Extra curricular course
[Biology/Medicine]

Course ID
cb-cbar

Class Start
Second week of lecture period

Timeline
Weekly

Estimated Effort
2 hours per week + 30 min. @home

Class Size
4 to 10 people

The RESEARCH COMMUNITY

This means we try to join learning and research as far as possible. We also encourage successful participants of the courses to consider projects in our field of research.

Lecture room: GPS Google maps link, 1.OG.

## COURSE STAFF

Prof. Korsching

## OBJECTIVE

Conscious and well-considered handling of scientific projects by computational methods (primarily R based).

## PREREQUISITES

None - but some interest in theoretical work and the software supporting such type of work is necessary.
For practice at home a notebook or PC is recommended.

## ABSTRACT

The field of computational biology is mainly based on programming languages, algorithms and applied mathematics. In this course we focus on the mathematical language R and try to learn step by step the philosophy of theoretical work and the art of computing including algorithms and mathematics/statistics.
Because this course mostly faces the situation that beginners and experienced students / scientists come together, we start, depending on the participants, the semester with the basics and evolve to advanced projects over time.
The reason for offering such an interdisciplinary and theoretical course for advanced students and post graduate researchers at the medical faculty, but also open for the natural science faculties, is based on the raising complexity in the life sciences and the increased use of theoretical methods. To bridge the existing gap between medical sciences and this highly theoretical field, the course is based on interactivity and a course level adjusted to the needs of the participants. Therefore we also encourage people with research interests in the experimental life sciences to join this course even if they to not expect to work on this field, to improve their insight into the methods and philosophy of bioinformatics

/ systems biology.

The audience is also encouraged to make their own proposals.

As per request we might extend to other processing languages too.

## Concept of the course

The course is designed for beginners and advanced users. Therefore the course starts with an introduction in a mixed course situation.

**CAVE:** The course is **not** a programmers only workshop
The following list is an overview on topics - concrete decisions are bound to the skills, proposals and the level of the participants

- Administrative details in the first double period

- Definition of computational biology / bioinformatics / tools / software etc.
- Review of all involved scientific and technical fields
- Purpose of the field
- Future prospective of the field

- Strategies and concepts in computational biology
- How to join theoretical concepts, measurement data and experimental methods
- Hands on example - projects / ideas of the audience

- Starting with an example: the general purpose mathematics platform - R
- The concepts behind this platform

- Some general notes on so called 'programming languages'
- The 'mathematical language' of R
- What is the gain to use R?  What is R not for?

- How to find resources?

- R basics - how to work?
    - notation, syntax, variables & data structures
    - control structures
    - decision structures
    - functions / modules / object orientated paradigms
    - R environments
    - R session file, R data files
    - the universe of packages and solutions
    - R repositories
    - linking to other languages / programs / computers
    - get an overview on the many graphics packages
    - 2D and 3D graphics

- Advanced R (selection)
    - graphic packages - graph, grid graphics, ggplot2
    - R color handling
    - debugging R code
    - Parallel processing in R
    - Optimizing R code
    - R and simulation
    - SQL database interface
    - creating documents with R
    - R web server

- Application scenarios (selection)
    - descriptive statistics
    - medical data analysis
    - survival analysis
    - ROC statistics
    - high throughput data analysis - e.g. microarray and NGS sequence data
    - machine learning
    - Bayes statistics

- Proposals from the audience

- **Retrospective seminar summary**
    - What did we learn from all these aspects?
    - How to generate a profit from all this theoretical knowledge for studies and practice in life science?