

# Bioinformatics and data analysis training module

Detailed information on courses and application are available at [the course website \(JOY\)](#).

**Summary:** This training provides excellent skills to search, process and utilize the available biomedical data in both scientific research and for applications in biotech companies. The training can be applied e.g. to identify potential target molecules or uncover biological pathways.

**Implementation:** Online self-study via Moodle. The teaching is given in English.

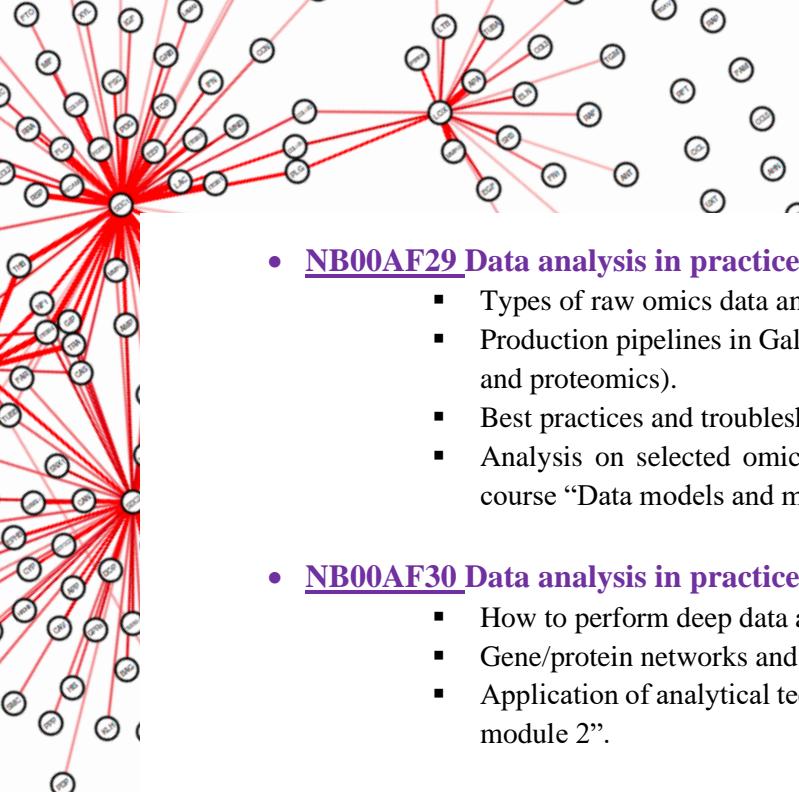
**Availability:** Courses are available from the beginning of the year 2022 until end of July in 2023. Courses can be taken anywhere and anytime enabling flexible completion.

**Eligibility:** Unemployed and employed persons with suitable background (biochemistry, biology, chemistry, mathematics, statistics, programming, medicine, dentistry etc.) and degree (Master's degree at University or University of Applied Sciences).

**How to apply:** Application of non-degree study rights via e-Forms. Application requires CV, degree certificate(s), short motivation letter (½ to 1 page, including e.g., previous work history, benefit of participation). Rolling application process is open until end of February in 2023.

## Description of courses in the training module:

- **NB00AF26 Introduction and learning tools to the continuous education in bioinformatics and data analysis, 1 ECTS**
  - Why are bioinformatics and data analysis important and how are they applied.
  - (Re-)Introduction to R programming and Moodle platform.
- **NB00AF27 Data models and methods, module 1, 3 ECTS**
  - Purposes and limitations of common machine learning (ML) methods.  
(e.g. supervised and unsupervised learning, linear regression and classification problems, decision tree learning, clustering algorithms).
  - Theory of ML model building (e.g. training/testing, standardization, cross-validation, overfitting, the curse of dimensionality, regularization).
  - How to build the model in practice (Preprocessing data, checking assumptions and implementation of a simple ML workflow in R)
- **NB00AF28 Data models and methods, module 2, 5 ECTS**
  - Types, purposes and limitations of common neural network (NN) models.
  - Theory of NN model building (e.g. activation functions, hidden layers, loss functions, batch sizes, back-propagation and learning rates).
  - Implement and train a simple neural network model in Keras using R and evaluate the model performance with Tensorboard.



- **NB00AF29 Data analysis in practice, module 1, 3 ECTS**
  - Types of raw omics data and their processing in Galaxy.
  - Production pipelines in Galaxy specific for data type (genomics, transcriptomics and proteomics).
  - Best practices and troubleshooting.
  - Analysis on selected omics data examples, applying the techniques from the course “Data models and methods, module 1”.
- **NB00AF30 Data analysis in practice, module 2, 3 ECTS**
  - How to perform deep data analysis, including inference approaches.
  - Gene/protein networks and integrated biological circuits.
  - Application of analytical techniques from the course “Data models and methods, module 2”.
- **NB00AF31-32 Final training project, parts 1 and 2, 5 ECTS each**

Topic and questions can be based on your own research interests or suggested by your employer.

- Evaluation of the analysis tasks and development of an analytical strategy.
- Methodological assessment and evaluation of the steps needed.
- Data analysis and review of the results, performance and aims.
- Further re-evaluate, improve and implement the improved analysis pipeline.
- Conclude the process and results from planning the analysis pipeline, its first implementation, improvement and final implementation.

### **Additional recommended courses:**

The missing background knowledge can be supplemented with additional recommended courses.

**Biosciences:** 744641S Bio-data (5 ECTS), 744640S Data mining and data-based models (5 ECTS)

Programming: 521141P Elementary programming (Ohjelmoinnin alkeet, 5 ECTS)

**Mathematics: 744643S Mathematics for biomedical data analysis (5 ECTS)**

## Further information:

# Emmi Alakärppä

## Education designer, JOY

emmi.alakarppa@oulu.fi

**Krista Juurikka**  
University teacher, FBMM  
krista.juurikka@oulu.fi