

Bio-IT, BioNT and Data Science

Supporting computational biology at EMBL



Renato Alves

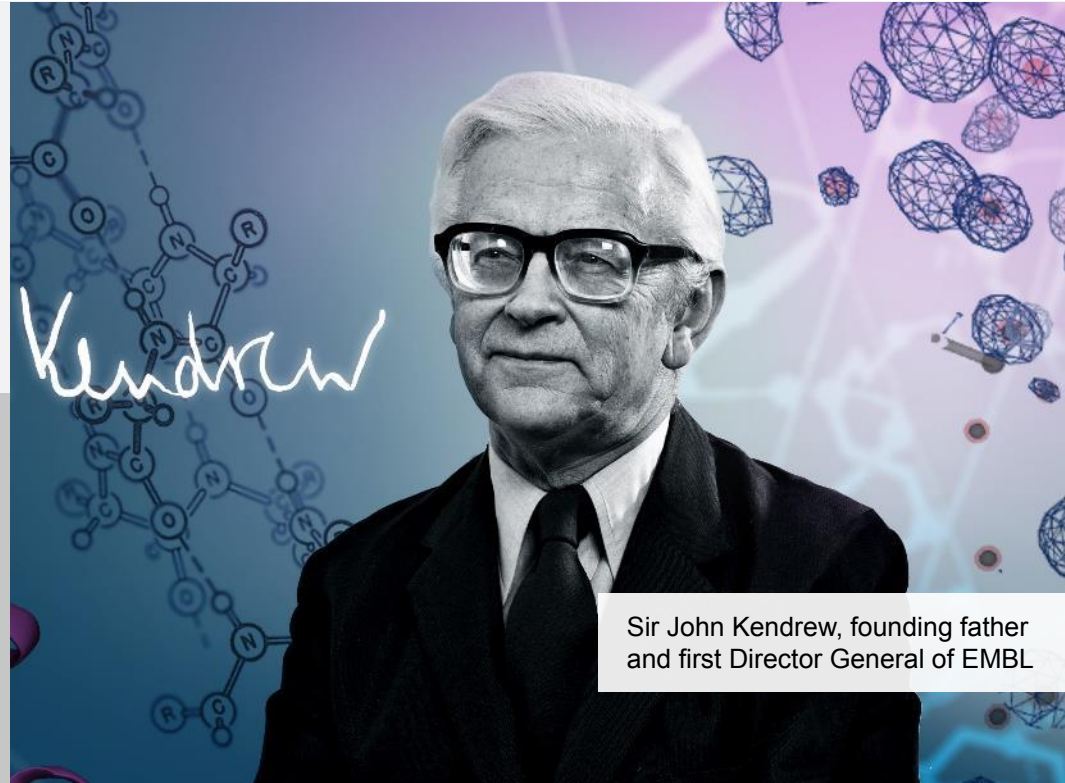
EIROforum - Workshop on knowledge management, 6th June 2024



What is EMBL?

**Europe's centre of excellence
in life science research,
services and training**

**Founded in 1974 by 10 states
(now 29 + 1 prospect) as an
intergovernmental organisation
to promote the molecular life
sciences in Europe and beyond.**



Sir John Kendrew, founding father
and first Director General of EMBL

EMBL sites – over 1700 people and more than 80 nationalities



EMBL-EBI
Bioinformatics



Grenoble
Structural biology



Barcelona
Tissue biology and
disease modelling



Hamburg
Structural biology



Heidelberg
Life sciences



Rome
Epigenetics
and neurobiology





European Molecular
Biology Laboratory

Knowledge exchange challenges

pan-EMBL activities

Hybrid/digital

Diverse needs

Multidisciplinarity

Rotation of personnel

Jargon

Overload of options



COMMUNITY



TRAINING



TOOLS



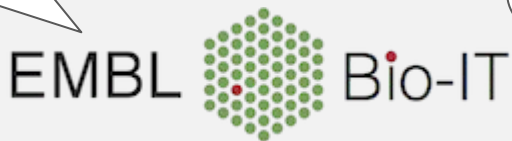
INFORMATION

The EMBL Bio-IT project



European Molecular
Biology Laboratory

Community initiative to build,
support and promote
computational biology
research at EMBL



The Project Managers:

Renato Alves



+ *BioNT project manager*
(Isabela Paredes Cisneros)

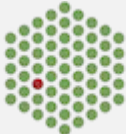
Lisanna Paladin



The EMBL Bio-IT project

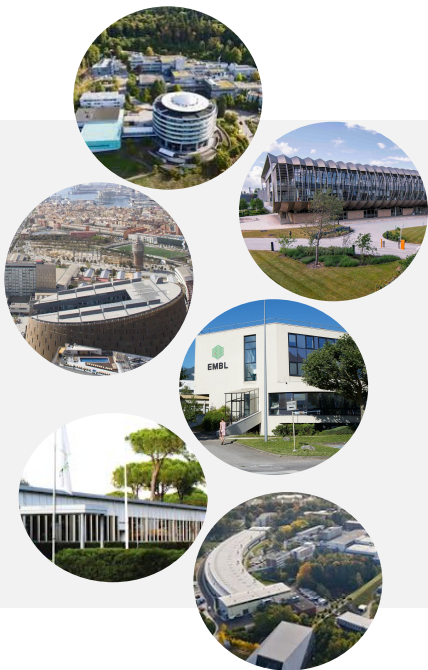
European Molecular
Biology Laboratory

Community initiative to build,
support and promote
computational biology
research at EMBL

EMBL  Bio-IT

#community





Data Science Centre

Managing, analysing and sharing data at a scale

Analysis Methods and Standards

- Artificial Intelligence
- Curated reference biodata sets
- Shared data analysis workflows

Integrated Data Management

- Data life cycles

Provision of Public Data Services

- Data resources
- EMBL tools



Technical Infrastructure

- Data processing compute infrastructure
- Network
- Storage

People and training

- Internal training and career development
- External training

Bio-IT website: information hub

bio-it.embl.de



Training

- Courses and events
- Course materials
- Consulting sessions

Resources

- EMBL chat
- GitLab
- Coding platforms
- Notes & survey tools
- Other resources

Community

- Coding clubs
- Grassroots
- Community blog

Information

- EMBL centers
- IT services info
- WikiHows
- Newcomers guide

The screenshot shows the bio-it.embl.de website interface. At the top, there is a navigation bar with links for Training, Community, Resources, Information, and Small Centers. Below the navigation bar, there is a main content area with several sections: Training, Community, Resources, and Information. Each section contains a list of links and brief descriptions of the resources available. For example, the Training section lists 'Upcoming courses', 'Course materials', 'Bio-IT meetings', 'Coding club', and 'Statistical consulting'. The Community section lists 'Get involved', 'Grassroots', 'Bio-IT community blog', 'EMBL', 'EMBL', 'EMBL', 'Bio-IT Beer Sessions', and 'Bio-IT Outreach & Travel'. The Resources section lists 'EMBL chat', 'GitLab', 'Coding platforms', 'Notes & survey tools', and 'Other resources'. The Information section lists 'EMBL centers', 'IT services info', 'WikiHows', and 'Newcomers guide'. On the right side of the page, there is a sidebar with a 'SHORTCUTS' section and a 'NEWSLETTER' section.

COMMUNITY

The World of Molecular Biology

Die Welt der Molekularbiologie

Textual content on the exhibition panel, including a diagram of a cell structure.

Small text on the right side of the exhibition panel, possibly a list of topics or credits.



Taskforce & community meetings

Hybrid format

CHANGED 19 HOURS AGO

Bio-IT Taskforce Meetings

Click here to join Zoom
Meeting ID: 963 1820 3177
Password: 604455

Online shared timer

Next taskforce meetings

- 2024-01-10 14:00
- 2024-03-06 14:00
- 2024-05-08 14:00
- 2024-07-03 14:00
- 2024-09-04 14:00
- 2024-11-06 14:00

Past taskforce meetings

- 2023 and prior years

2024-01-10 14:00 (German-time) - hybrid

Agenda

- Introductions & newcomers
- The Lean Coffee model

Todo from past meetings

Expand list

Attending

Please add your name here.
Create new bullet points if necessary:

-
-
-
-

CHANGED A FEW SECONDS AGO

Data Science Community Meeting

click here to join the zoom meeting

Meeting ID: 823 807 3039
Passcode: r9V9NE

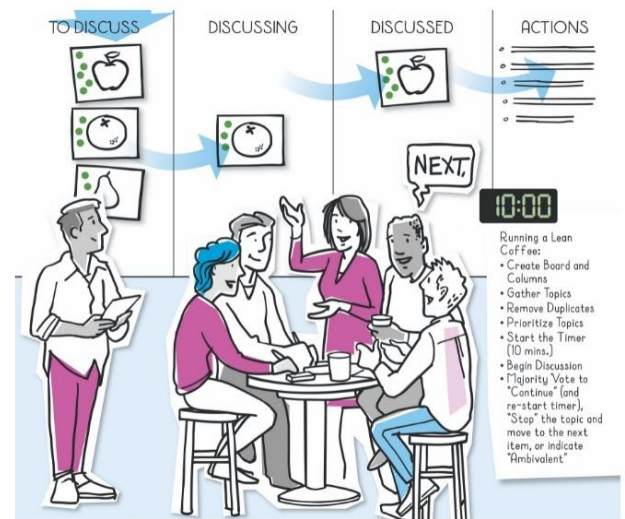
Next DS community meetings

- 2024-02-07 14:00
- 2024-04-03 14:00
- 2024-06-05 14:00
- 2024-08-07 14:00
- 2024-10-02 14:00
- 2024-12-04 14:00

Past DS community meetings

- 2023-12-06 14:00
- 2023-10-04 14:00

Lean Coffee model



Coding clubs

Hybrid format

EPUG	A bi-weekly meeting for Python users at EMBL.
emblr	A bi-weekly meeting for R users at EMBL.



python™



plotly

Programming Clubs: EPUG and emblr

- Every Tuesday, 15:00 Germany time
- Alternating Python / R
- Introduce new topics?
 - Image analysis

Workflow Management Systems Club New

- Monthly, first Thursday, 10:00 Germany time

BioinfoRome



- Flexible schedule, discussion about technical questions

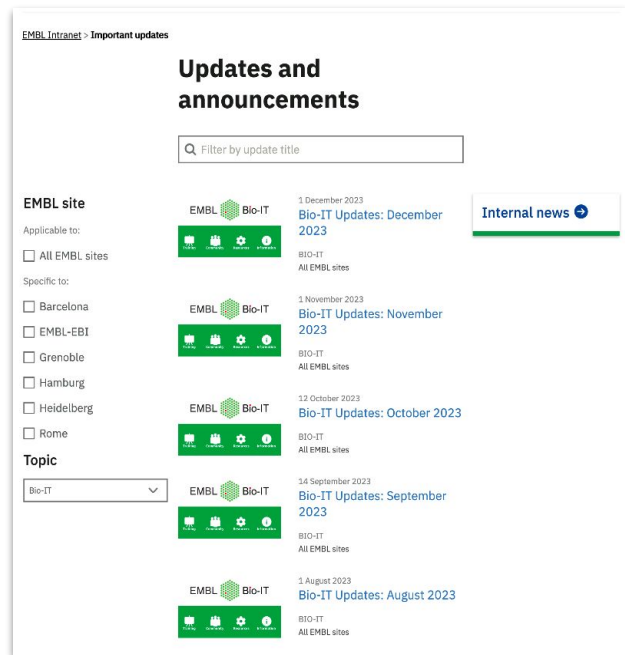
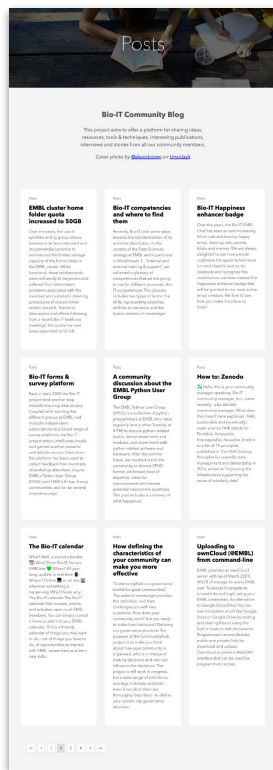
Community blog and Newsletter

grp-bio-it.embl-community.io/blog

embl.org/internal-information

Bio-IT Community Blog

This project aims to offer a platform for sharing ideas, resources, tools & techniques, interesting publications, interviews and stories from all our community members.



TRAINING



Courses and events

BIO-IT ▾ TRAINING ▾ COMMUNITY ▾ RESOURCES ▾ INFORMATION ▾ EMBL CENTRES ▾

UPCOMING COURSES

COURSE MATERIALS

ONLINE LEARNING
RESOURCES

TRAINING INFORMATION

Upcoming Courses

List of course and events that are planned, date/time and place (in person / hybrid / online)

Special mention: [Bio-IT and Data Science Taskforce Meetings](#)

Training plan 2024

List of courses that we plan to organise, including estimated dates, status, links to materials/notes, lead and team

Invite to express training needs through the [Bio-IT catalogue](#)



Hybrid format

VPN only

Courses and events

catalogue.bio-it.embl.de

Bio-IT Training Catalogue Skills & Topics Statistics

Bio-IT Portal

Training Catalogue

This site hosts a catalogue of the past training events offered by Bio-IT.

Browse recent events or use the **search** feature to find events according to your desired search parameters. Results will populate the table below. Click on a Course or Event title to learn more.

Search Catalogue

Keyword

Trainers

Start Date

-

Experience Level

Location

- EMBL Hamburg
- EMBL Barcelona
- EMBL-EBI Hinxton
- EMBL Heidelberg
- EMBL Rome
- EMBL Grenoble
- Online
- IPMB - Heidelberg University

Internal/External

Event Format

Skills

- Benchmarking bioinformatics tools
- Biological modelling
- Biological networks analysis
- Cloud computing
- Cluster computing (HPC)
- Command-line computing
- Computational workflow management
- Data management and curation
- Exploratory data analysis and visualisation (EDA)
- GPU computing
- Image analysis
- Programming languages
- Software project management
- Statistics and machine learning
- System administration
- Text mining
- Using specialised research software
- Web technologies

Topics

- Bioc mining
- Cancer genomics and personalised medicine
- Genomics and comparative genomics
- Instructor training
- Metabolomics
- Metagenomics and other meta-omics
- Multi-omics and single-cell integrative analysis
- Neurobiology and neuroinformatics
- Open Science
- Proteomics and protein analysis
- Structural biology

Course: Version Control with Git

Course Overview

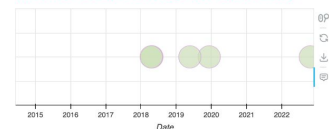
Description

Do you struggle to keep track of the changing versions of your scripts/programs? Have you ever made changes to a file, only to wish shortly afterward that you could easily get back the old version? Do you want to learn how to harness the awesome features of GitLab for collaboration, issue tracking, software publication, etc? This course aims to teach good practice in version control, using git in conjunction with the EMBL GitLab system. Participants will learn: how to maintain a history of changes with Git and how to compare differences between versions; how to restore old versions of files; how to synchronise local versions with a remote repository; how multiple developers can collaborate effectively on a project, and how to resolve conflicting changes; how to manage a project through the GitLab interface; how to choose the right license for a project; how to work with Git directly within your development environment (e.g. RStudio). The course will provide an introduction to Git via three different interfaces: the command line, a graphical client, and GitLab.

By the Numbers

Events Offered	5
Course Attendance	Minimum 11
	Maximum 37
	Total 124
Course Duration (Days)	Minimum 1
	Maximum 2
	Total 6

Timeline of Course Offerings (2018 to Present)



Latest Course Resources:

Bio-IT Record (2022-10-24):
<https://bio-it.embl.de/events/version-control-with-git-git-embl-de-git-week-basic-module/>

Git Repository (2022-10-24):
<https://github.com/swcarpentry/git-novice>

Reference materials (2022-10-24):
<https://pad.bio-it.embl.de/sd2Mk8YVSbuMg3KI-nL6yA>

Events Offered

Version Control with Git & git.embl.de | Git week, basic module

Bio-IT Record: <https://bio-it.embl.de/events/version-control-with-git-git-embl-de-git-week-basic-module/>

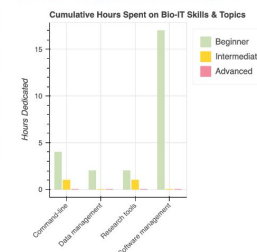
Git materials: <https://github.com/swcarpentry/git-novice>

Reference materials: <https://pad.bio-it.embl.de/sd2Mk8YVSbuMg3KI-nL6yA>

Start Date	2022-10-24
Course Duration (Days)	2
Course Attendance	27
Instructor(s)	Lisanna Paladin, Renato Alves
Location(s)	EMBL Heidelberg Online
Reach	Internal
Course Level	Beginner
Hours per Skill	Beginner Command-line computing: 3 Beginner Software project management: 3 Beginner Using specialised research software: 2

VPN only

Curious how your time will be spent?



Interested in attending this course?

Click the button below to indicate your interest and join the mailing list for upcoming instances of the course.

Consortium project

Partners: 6 academic, 3 SME, 4 countries



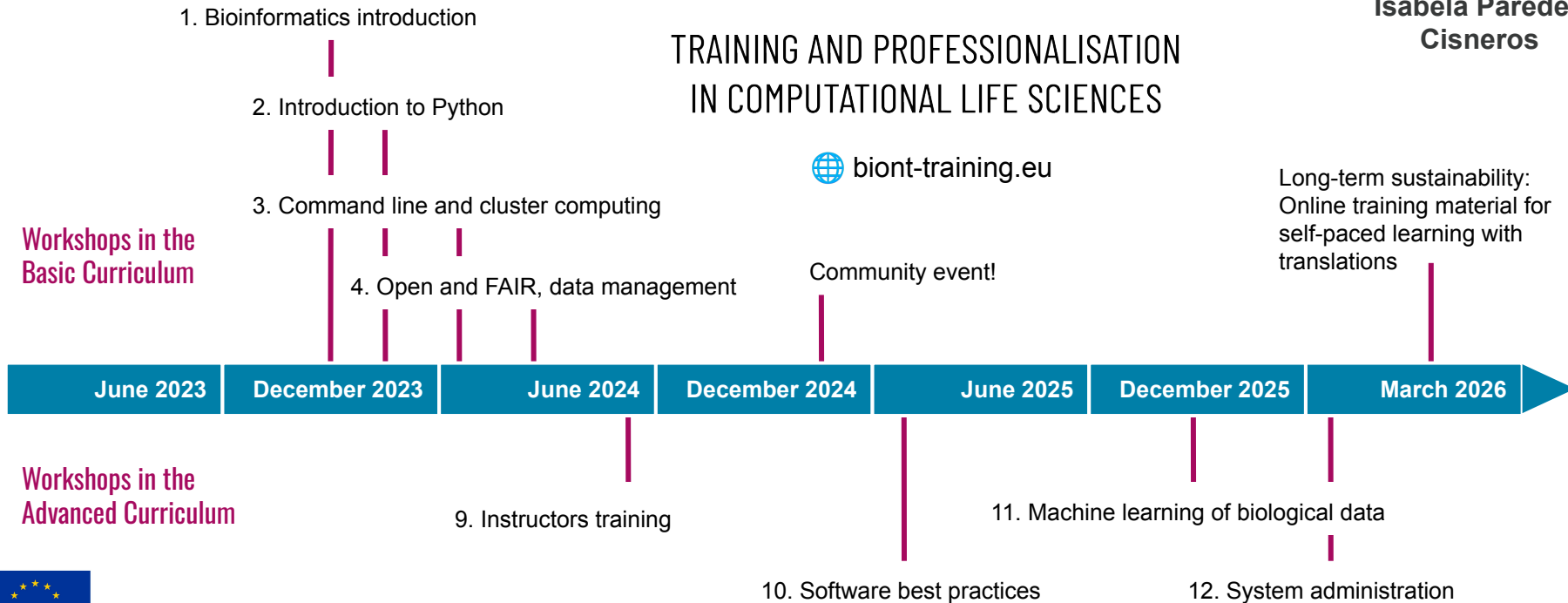
Coordinated by:

**Isabela Paredes
Cisneros**

TRAINING AND PROFESSIONALISATION IN COMPUTATIONAL LIFE SCIENCES

 biont-training.eu

Long-term sustainability:
Online training material for
self-paced learning with
translations



Workshops in the
Basic Curriculum

Workshops in the
Advanced Curriculum



Co-funded by
the European Union



Course materials

Course Materials

Links to training materials, in different formats when available and contact of the main maintainer.

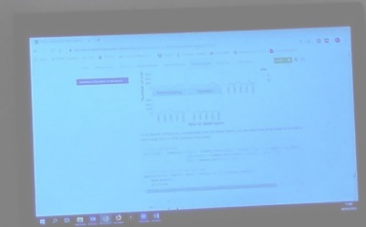
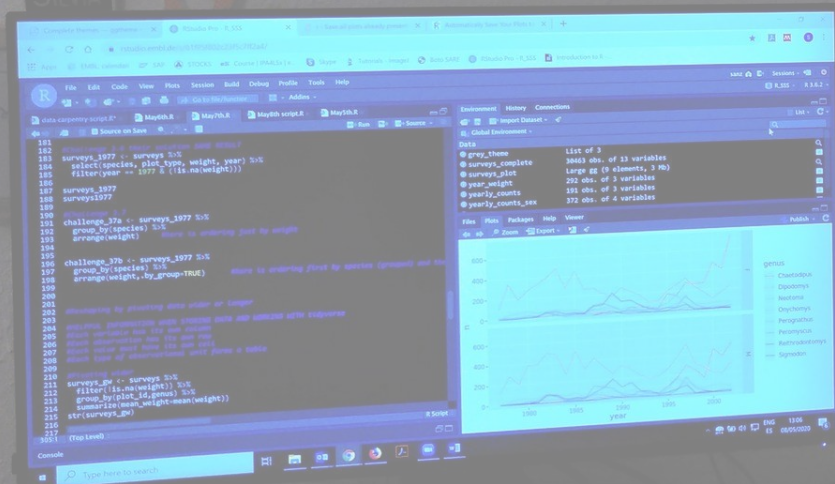
Sections:

- Bio-IT courses
- Cluster training
- Singularity training
- CBNA courses
- CBM courses
- Predoc course



Recommended Resources for Online Learning

INFRASTRUCTURE





Welcomebot BOT 3:06 PM

Awesome! I've added you to the following channels:

- Bio-IT 🚀 Bio-IT channel for general bioinformatics, computational support
- Bio-IT-announce 🚀 Bio-IT channel for course/workshop and Bio-IT activity announcements
- cluster 🚀 EMBL Heidelberg HPC cluster support
- GBCS 🚀 Genome Biology Computational Support channel
- Mathematical Modelling 🚀 Genome Biology Computational Support channel
- STOCKS 🚀 Support channel for STOCKS EMBL's -GBCS electronic lab and data management system

You may also be interested in the following (click their name to join):

- Folding / AlphaFold 🚀 A channel to discuss protein folding with AlphaFold and related technologies
- How could be done better 🚀 A channel to discuss scientific figures and other kinds of scientific output in a constructive way

Awesome! I've added you to the following discussion channels for different programming related technologies:

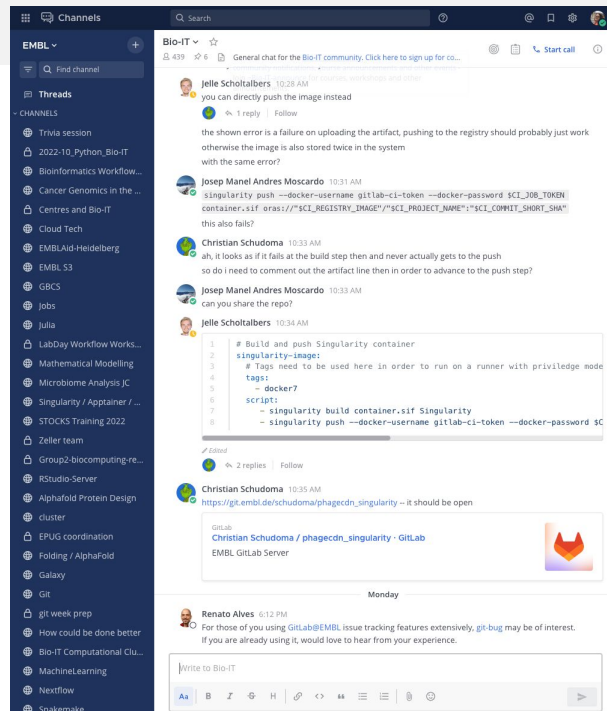
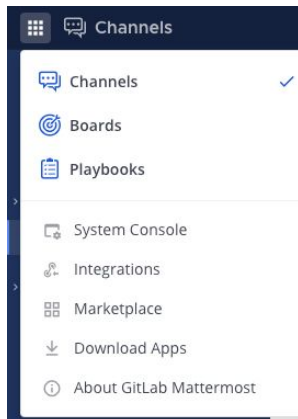
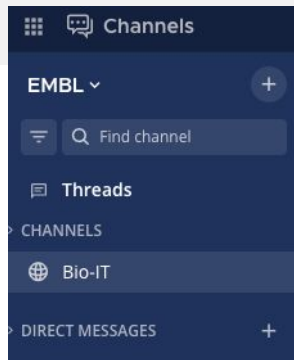
- Git 🚀 Git and EMBL's GitLab
- Julia 🚀 Julia
- MachineLearning 🚀 For classic and deep learning discussions
- Python 🚀 Python
- R_and_statistics 🚀 R and statistics
- RStudio-Server 🚀 RStudio EMBL server (by -GBCS)

Awesome! I've added you to the following support channels:

- Galaxy 🚀 Galaxy@EMBL (by -GBCS)
- Nextflow 🚀 Nextflow workflow manager
- Snakemake 🚀 Snakemake workflow manager
- Singularity / Apptainer / Docker 🚀 Singularity containers

Awesome! I've added you to the following channels:

- Jobs 🚀 Announce jobs and other opportunities in or outside of EMBL.
- marketplace 🚀 General trading, buy&sell, giveaways, lost&found, ...
- EMBLAid-Heidelberg 🚀 EMBLAid channel for Heidelberg (see also channels for other sites).

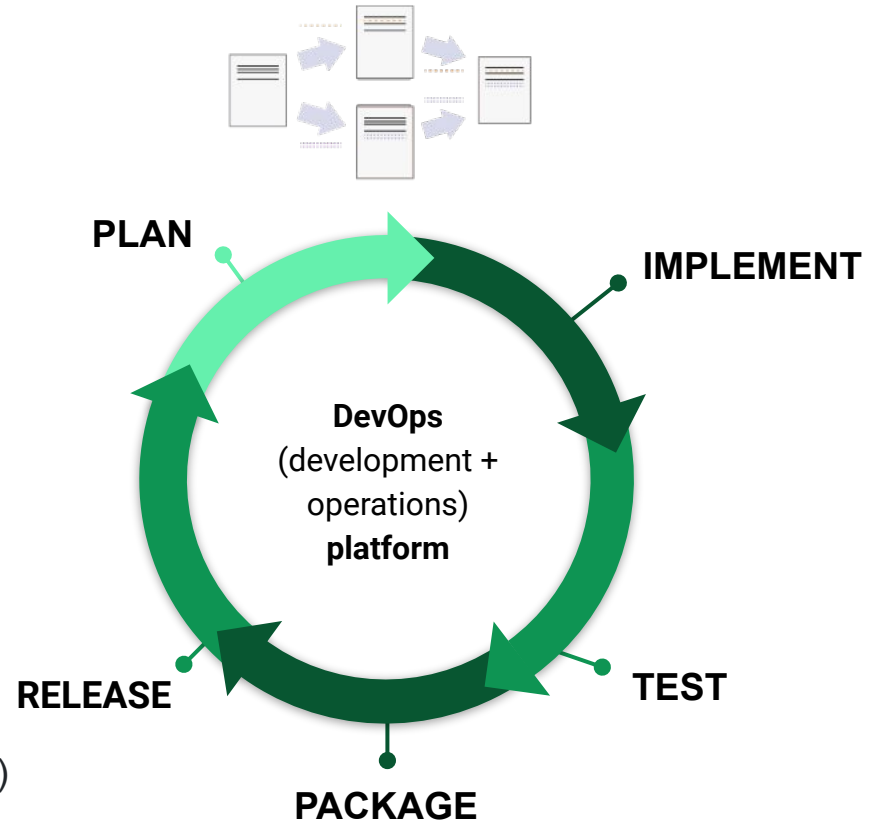




git.embl.de

Git = Version control system

- System that records changes made to a project (group of files) over time
 - Keeps versions in a smart way
 - Allows reverting back to a previous state
- Essential to (programming) projects involving several people
 - Allows branching (working separately on different features)
 - Allows to compare/merge changes made
 - Allows synchronization of repositories in different places (distributed version control)

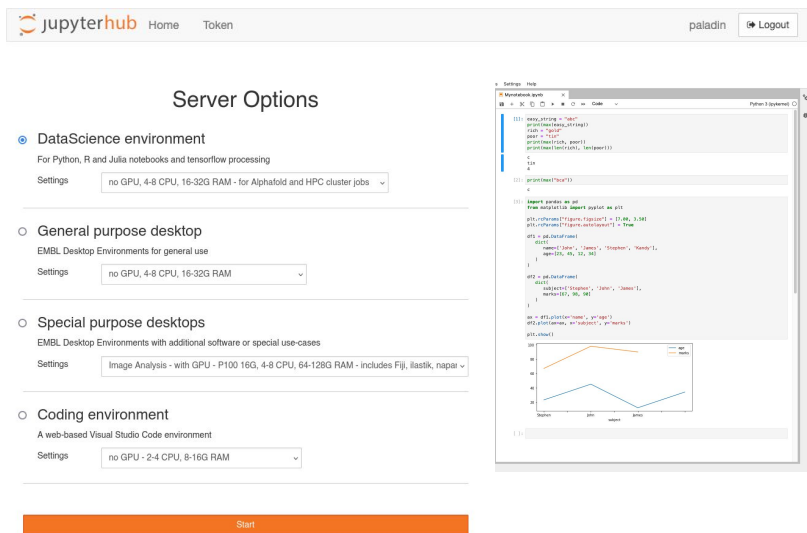


⚙️ Coding platforms



jupyterhub.embl.de

 **jupyterhub**



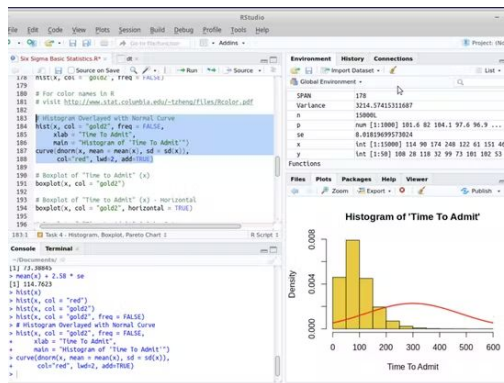
Server Options

- DataScience environment**
For Python, R and Julia notebooks and tensorflow processing
Settings: no GPU, 4-8 CPU, 16-32G RAM - for AlphaFold and HPC cluster jobs
- General purpose desktop**
EMBL Desktop Environments for general use
Settings: no GPU, 4-8 CPU, 16-32G RAM
- Special purpose desktops**
EMBL Desktop Environments with additional software or special use-cases
Settings: Image Analysis - with GPU - P100 16G, 4-8 CPU, 64-128G RAM - includes Fiji, Ilastik, napari
- Coding environment**
A web-based Visual Studio Code environment
Settings: no GPU - 2-4 CPU, 8-16G RAM

Start



Genome Biology Computational Support



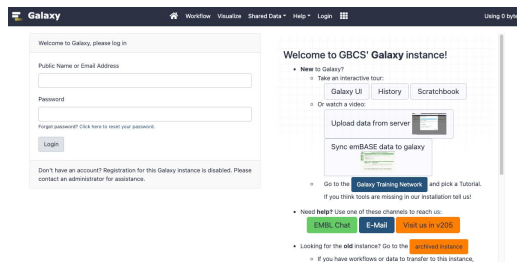
```
## Histogram overlaid with Normal curve
hist(x, col = "gold", freq = FALSE,
     xlab = "Time to Admit",
     main = "Histogram of Time To Admit",
     col="red", lwd=2, add=TRUE)

## Resplot of "Time to Admit" (x)
plotlet(x, col = "gold")

## Resplot of "Time to Admit" [1] horizontal
plotlet(x, col = "gold", horizontal = TRUE)
```



rstudio.embl.de



Welcome to Galaxy, please log in

Public Name or Email Address:

Password:

Forgot password? Click here to reset your password.

Log in

Don't have an account? Registration for the Galaxy instance is disabled. Please contact an administrator for assistance.

Welcome to GBCS' Galaxy instance!

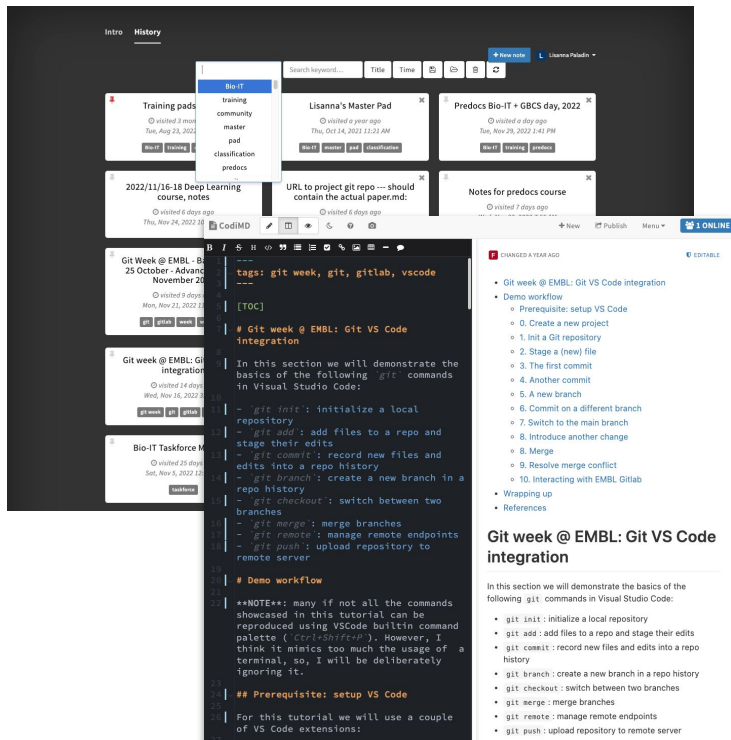
- New to Galaxy?
 - Take an interactive tour:
 - Galaxy UI
 - Or watch a video:
 - Upload data from server
 - Sync emBASE data to galaxy
- Go to the [Galaxy Training Network](#) and click a Tutorial. If you think books are missing in our installation test set.
- Need help? Use one of these channels to reach us:
 - EMBL Chat
 - E-Mail
 - Visit us in 2020
- Looking for the old instance? Go to the [previous instance](#).
- If you have workflows or data to transfer to this instance,



galaxy.embl.de

⚙️ Notes/Presentation and survey tools

pad.bio-it.embl.de  CodiMD



The screenshot shows the CodiMD interface with a grid of training pads. One pad is titled "Git week @ EMBL: Git VS Code integration". The code editor shows the following content:

```

tags: git week, git, gitlab, vscode
---
[TOC]

# Git week @ EMBL: Git VS Code integration

In this section we will demonstrate the basics of the following git commands in Visual Studio Code:

- git init: initialize a local repository
- git add: add files to a repo and stage their edits
- git commit: record new files and edits into a repo history
- git branch: create a new branch in a repo history
- git checkout: switch between two branches
- git merge: merge branches
- git remote: manage remote endpoints
- git push: upload repository to remote server

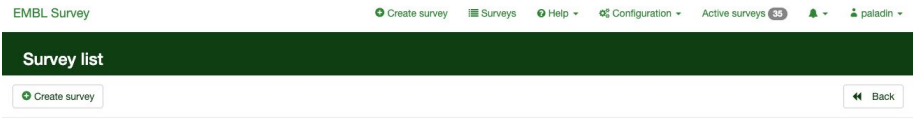
## Demo workflow

**NOTE**+: many if not all the commands showcased in this tutorial can be reproduced using VSCode builtin command palette (Ctrl+Shift+P). However, I think it mimics too much the usage of a terminal, so, I will be deliberately ignoring it.

## Prerequisite: setup VS Code

For this tutorial we will use a couple of VS Code extensions:
    
```

survey.bio-it.embl.de 

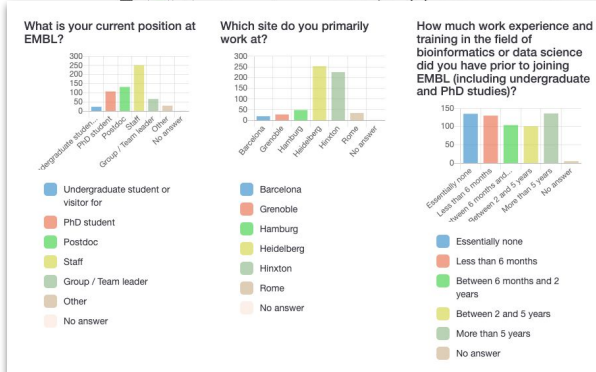


The screenshot shows the EMBL Survey dashboard. At the top, there are navigation links: "EMBL Survey", "Create survey", "Surveys", "Help", "Configuration", "Active surveys (35)", and "paladin". Below this is a "Survey list" section with a "Create survey" button and a "Back" button.

Survey list | Survey groups

Search: Status: (Any) Group: (Any group) Search

Action	Survey ID	Status	Title	Group	Created	Owner	Anonymized responses	Partial	Full	Total	Closed group
<input type="checkbox"/>	646227	▶	Deep Learning course - Day 2 feedback	Default	15.11.2022	blot	No	0	4	4	No
<input type="checkbox"/>	374587	▶	Deep Learning course - Day 1 feedback	Default	15.11.2022	blot	No	4	5	9	No
<input type="checkbox"/>	764775	▶	Post-workshop survey Python course	Default	28.10.2022	blot	No				<input type="button" value="Close"/>
<input type="checkbox"/>	781197	▶	Pre-workshop survey Git course	Default	25.10.2022						
<input type="checkbox"/>	412597	▶	Pre-workshop survey Python course	Default	20.10.2022						



What is your current position at EMBL?

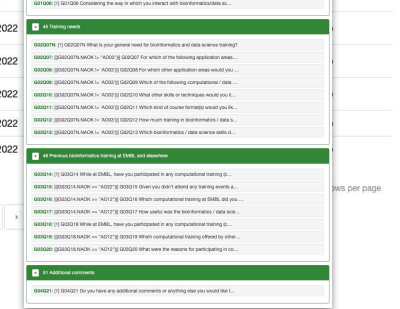
Category	Undergraduate student or visitor for	PhD student	Postdoc	Staff	Group / Team leader	Other	No answer
Barcelona	~100	~100	~100	~100	~100	~100	~100
Grenoble	~100	~100	~100	~100	~100	~100	~100
Hamburg	~100	~100	~100	~100	~100	~100	~100
Heidelberg	~100	~100	~100	~100	~100	~100	~100
Hinxton	~100	~100	~100	~100	~100	~100	~100
Rome	~100	~100	~100	~100	~100	~100	~100
No answer	~100	~100	~100	~100	~100	~100	~100

What site do you primarily work at?

Category	Barcelona	Grenoble	Hamburg	Heidelberg	Hinxton	Rome	No answer
Barcelona	~100	~100	~100	~100	~100	~100	~100
Grenoble	~100	~100	~100	~100	~100	~100	~100
Hamburg	~100	~100	~100	~100	~100	~100	~100
Heidelberg	~100	~100	~100	~100	~100	~100	~100
Hinxton	~100	~100	~100	~100	~100	~100	~100
Rome	~100	~100	~100	~100	~100	~100	~100
No answer	~100	~100	~100	~100	~100	~100	~100

How much work experience and training in the field of bioinformatics or data science did you have prior to joining EMBL (including undergraduate and PhD studies)?

Category	Essentially none	Less than 6 months	Between 6 months and 2 years	More than 2 and 5 years	More than 5 years	No answer
Barcelona	~100	~100	~100	~100	~100	~100
Grenoble	~100	~100	~100	~100	~100	~100
Hamburg	~100	~100	~100	~100	~100	~100
Heidelberg	~100	~100	~100	~100	~100	~100
Hinxton	~100	~100	~100	~100	~100	~100
Rome	~100	~100	~100	~100	~100	~100
No answer	~100	~100	~100	~100	~100	~100



EMBL Training week

0101001 (1) 0101001 What is your general need for bioinformatics and data science training?

0101002 (1) 0101002 Which of the following application areas...

0101003 (1) 0101003 For which of the following application areas would you...

0101004 (1) 0101004 Which of the following application areas would you...

0101005 (1) 0101005 Which of the following application areas would you...

0101006 (1) 0101006 Which of the following application areas would you...

0101007 (1) 0101007 Which of the following application areas would you...

0101008 (1) 0101008 How much of your work time do you typically spend on providing traini...

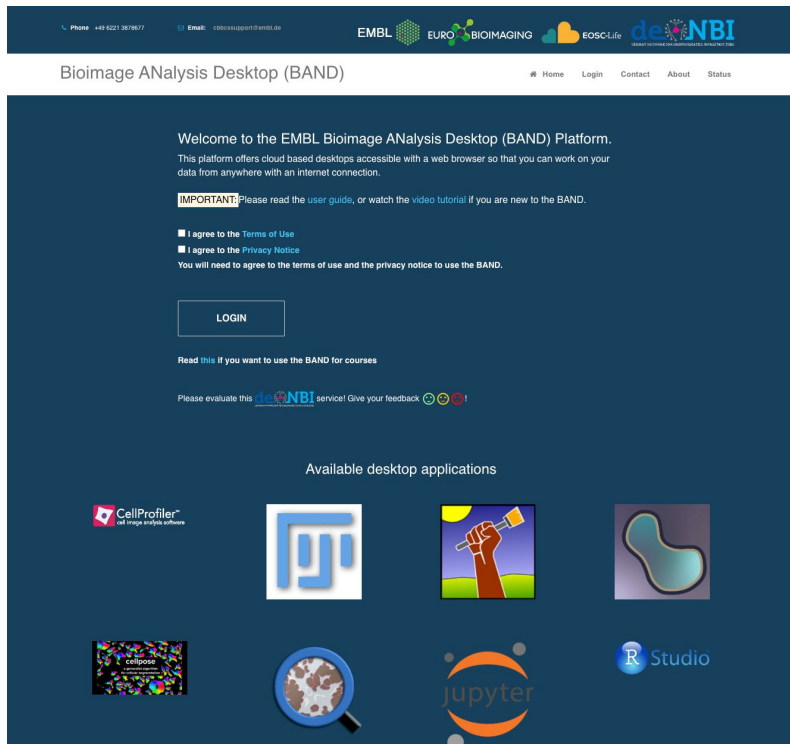
0101009 (1) 0101009 How much of your work time do you typically spend on providing traini...

0101010 (1) 0101010 Considering the way in which you interact with bioinformatics traini...

EMBL 50 YEARS 1974-2024

Other resources

band.embl.de



Phone +49 6221 387617 | Email cbcsupport@embl.de

EMBL EURO BIOIMAGING EOSC4Life www.embl.de www.eosc4life.eu www.nbi.ac.uk

Bioimage Analysis Desktop (BAND)

Home Login Contact About Status

Welcome to the EMBL Bioimage Analysis Desktop (BAND) Platform.
This platform offers cloud based desktops accessible with a web browser so that you can work on your data from anywhere with an internet connection.

IMPORTANT: Please read the [user guide](#), or watch the [video tutorial](#) if you are new to the BAND.

- I agree to the Terms of Use
- I agree to the Privacy Notices

You will need to agree to the terms of use and the privacy notice to use the BAND.

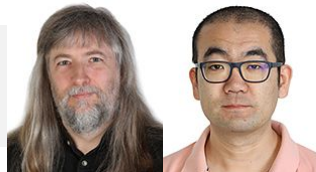
LOGIN

Read [this](#) if you want to use the BAND for courses

Please evaluate this [EMBL NBI](#) service! Give your feedback

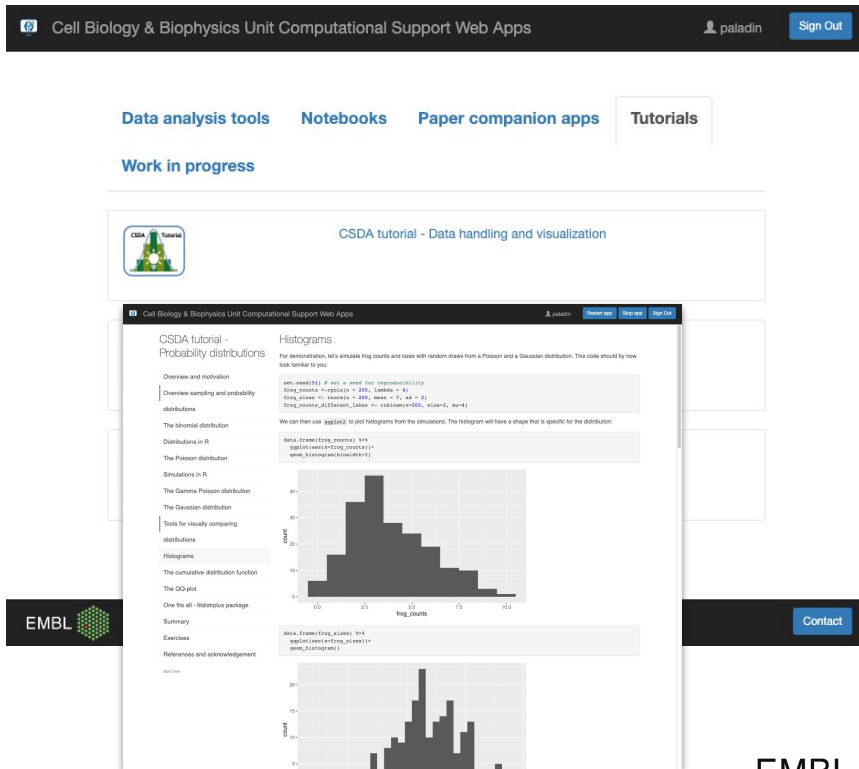
Available desktop applications

- CellProfiler™ cell image analysis software
- ImageJ
- Microscopy Image Analysis (MIA)
- ImageLab
- ImageM
- Jupyter
- R Studio



Cell Biology and Biophysics Computational Support

cbcs-01.embl.de/shinyapps/



Cell Biology & Biophysics Unit Computational Support Web Apps | paladin | Sign Out

Data analysis tools | Notebooks | Paper companion apps | Tutorials

Work in progress

CSDA tutorial - Data handling and visualization

CSDA Tutorial

CSDA tutorial - Probability distributions

Overview and motivation

Overview sampling and probability distributions

The binomial distribution

Distributions in R

The Poisson distribution

Simulators in R

The Gamma Poisson distribution

The Gaussian distribution

Tools for visually comparing distributions

Histograms

The cumulative distribution function

The QQ-plot

One fit all - Histogram package

Summary

Exercises

References and acknowledgment

```
## Simulate log counts and sizes with random draws from a Poisson and a Gaussian distribution. This code should be new to you (but not to me).
set.seed(5) # set a seed for reproducibility
freq_counts <- rpois(n = 100, lambda = 4)
freq_sizes <- rnorm(n = 100, mean = 7, sd = 3)
freq_counts_diff <- rabinom(n=100, size=2, m=4)
```

We can then use `ggplot2` to plot histograms from the simulations. The histogram will have a shape that is specific for the distribution:

```
data <- frame(freq_counts, size)
ggplot(data, aes(freq_counts)) +
  geom_histogram(binwidth = 5)
```


EMBL



Contact


⚙ Other resources



ds-wizard.embl.org

 DS Wizard

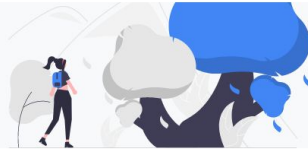
-  Dashboard
-  Projects

 Renato Alves
Researcher

⏪ Collapse sidebar

Welcome, Renato!

As a researcher, you create and collaborate on data management plans.



You have no projects yet, start by creating some.

→

Create Project

Project is a workspace where you create your DMP. It is based on a knowledge model, which contains knowledge about what should be asked and how based on the research field or organization's needs. You can use document templates to transform the answers into a document. This document can be anything, from PDF to machine-actionable JSON.

You can create a new project from a project template that data stewards prepare for you to have an easier start or from scratch where you set up everything yourself.

[Create](#)

P265

Community building and training in the EMBL Bio-IT project

Renato Alves, Lisanna Paladin

THE CONTEXT

- Interdisciplinary organization
- 6 sites in 6 EU member states
- Dynamic research content
- Diverse scientific goals
- Highly collaborative environment
- Planetary scale science

THE CHALLENGE

- Remote work - physical pandemic
- Distributed nature of the workforce
- Scattered and voluntary expert knowledge
- Standardization of training process
- Broad range of topics and needs
- Advancement and delivery
- Data privacy, volume and computation

THE SOLUTION

TRAINING

- Focus on learning objectives, content to learn to do
- Flexible, modular format with increasing complexity
- UK skills and IT topics covered
- Focus on preparing for data measurement

INFORMATION

- Built website as main hub of information
- Document guide for accelerated delivery plan
- Built long website for web for the community
- Remote guidance from content to announcements

TOOLS

- Website to guide for software and project management
- Automated for software test and code comparison
- Automated data analysis pipeline
- Automated data analysis pipeline

COMMUNITY

- Regularly scheduled meetings and webinars
- Virtual Bio-IT Academy (2020)
- Virtual Bio-IT Academy (2021)
- Virtual Bio-IT Academy (2022)
- Virtual Bio-IT Academy (2023)

bio-it@embl.de
 bio-it@embl.de
 @EMBL_BioIT

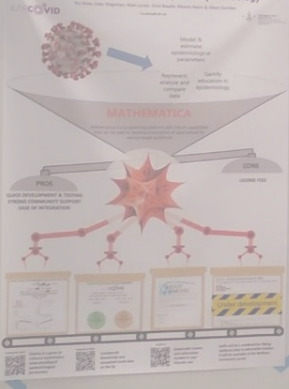
EMBL Bio-IT de-NBI

TRAINING TRACK

EMBL-EBI

P263

Mathematica as a platform to develop tools for education, data representation, and analysis in epidemiology



Clear: Improving reproducibility and datasets based on gene identifiers



INFORMATION

Consulting sessions

Bio-IT Drop-In Sessions

Drop-in sessions with [Renato Alves](#) and [Lisanna Paladin](#) run **10:00-12:00 every Tuesday** in the cafeteria/staff lounge at EMBL Heidelberg and/or online via Zoom. Check the individual listings on [the online events calendar](#) for Zoom connection details.

For advice & help with:

- General bioinformatics
- Command Line
- Scripting
- Task/analysis automation
- Biocomputing at EMBL Heidelberg



Newcomers guide

Bio-IT Newcomers Guide

Note that several linked pages in this document can be accessed only from inside the EMBL network

Table of contents

- [About Bio-IT Project](#)
- [Members of the Bio-IT Community](#)
- [Computational Resources](#)
- [Technical Setup](#)
- [Events](#)
- [Contacts](#)

About this document

This document is developed to help newcomers to learn all about computational systems at EMBL-HD, Bio-IT, and (almost) everything that will help them intergrate with the bio-computational community of EMBL Heidelberg.

The information in this guide overlaps with the material contained in the other documents useful for newcomers to EMBL. We also recommend that you check out the following resources:

- [Bio-IT EMBL website](#) for general information about Bio-IT activities.
- [EMBL website](#) for general information about EMBL.
- [EMBL newcomer's booklet](#): including information about first day best practices, housing, banking, medical service, childcare and family care, transportation, free time activities, additional informaton about Germany.
- [EMBL Staff Association website](#), including SA events calendar, forums and working groups and initiatives for the EMBL-ites' well being.
- [Heidelberg city guide](#), featuring city tour and events announcements.
- [Szilard library guide](#), including the catalogue and access information.

If you can't find the information that you were looking for, please go ahead and directly interact with the EMBL-ites via the [EMBL chat](#).

Acknowledgments

This guide borrows heavily (with permission) from the newcomers guides previously developed in the Zaugg and Huber Labs at EMBL. We thank the authors of those documents for their efforts and help.

Several sections also link out to the excellent [Netherlands eScience Center Guide](#).

Table of contents

- [About Bio-IT Project](#)
- [Members of the Bio-IT Community](#)
- [Computational Resources](#)
- [Technical Setup](#)
- [Events](#)
- [Contacts](#)

EMBL centers

Centre for Biomolecular Network Analysis

The mission of the Centre is to disseminate expertise, know-how and guidance in the field of biological network integration and analysis throughout computational biologists and experimentalists alike. For expert users embedded in experimental groups and spread around the Research Units, it will serve as a platform to share resources, exchange and learn about advances and new approaches in the computational networks. At the same time the Centre also offers support to talents with less computational expertise, who wish to pursue large scale or to place small-scale mechanistic experiments in the context of networks. The education and training of EMBL researchers in the field of data integration and network analysis is a top priority of the Centre.

Centre for Bioimage Analysis

The Centre for Bioimage Analysis (CBA) supports scientists in extracting quantitative information from images acquired with light- or electron-microscopy.

Support is provided on various aspects of image analysis, such as

- image data handling
- quantitative image interpretation and visualisation
- image registration
- object segmentation and tracking
- morphological and intensity quantifications.

The CBA closely collaborates with the Advanced Light Microscopy Facility (ALMF) and the Electron Microscopy Core Facility (EMCF) in order to ensure optimal image acquisition modalities for downstream quantitative analysis.

Together with ALMF, EMCF and EMBL's IT department



Centre for Biological Modelling

The mission of the Centre for Biological Modelling (CBM) is to support interested EMBL scientists in adopting mathematical modelling techniques, and to foster the interchange of existing modelling expertise throughout EMBL. The services offered by the CBM are open to EMBL scientists of all career levels. Just write me an email or drop by at the Centres Office (Room 118 in Building 1E).

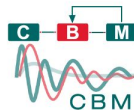
Please use the menu in the right hand side column of the page (scroll down to EMBL CENTRES) to explore the CBM and the different resources.

The CBM provides consultation appointments, performs collaborative modelling tasks, trains EMBL members in relevant programming languages and software tools and will foster collaboration and interactions between EMBL researchers using seminars, journal clubs as well as interdisciplinary retreats.

You have the data and the in depth knowledge of the system you are investigating and you know which question you want to answer. The CBM can provide you with guidance on modelling approaches and tools to support your attempts to answer your questions. Typically questions that can be tackled by modelling are:

- Is your current understanding of the systems mechanisms in accordance with your data?
- If not, what additional mechanism could improve consistency between data and what you believe is the process that leads to these data?
- Given your data, which of your hypotheses on how the components of your system interact is the most likely one?
- What is a quantitative measure of the effect of your experimental perturbations on your read out?
- Given your knowledge of the system you are studying, are there experiments that make more sense than others?
- Are there any feedback interactions in your system which are not obvious from the schematic representation?

The current expertise of the centre resides in first principles modelling using Ordinary, Random and Stochastic Differential Equations, Constraint-Based Modelling such as Flux Balance Analysis as well as Boolean and Mixture Modelling approaches. Nevertheless, inquiries concerning any kind of modelling are always welcome.



Bio-IT

The Bio-IT community initiative provides support to computational biology research through training, courses, consulting, networking opportunities, and computational resources and tools.

Contacts: Renato Alves and Lisanna Paladin (bio-it@embl.de)
Stop by our Drop-In Sessions Tuesdays 10am-12pm



Centre for Bioimage Analysis (CBA)

The CBA supports you in extracting quantitative information from images acquired with light- or electron-microscopy.

Contacts: Christian Tischer, Sebastian Gonzalez, and Arif Khan (image-analysis-support@embl.de)
Send a mail to schedule a consultancy session.



Centre for Biological Modelling (CBM)

The CBM aims to support you to adopt mathematical modeling techniques into your everyday research.

Contact: Eva Geissen (eva.geissen@embl.de)
Write me an email to request an appointment or come over to Room 118.



Centre for Statistical Data Analysis (CSDA)

The CSDA helps you to use adequate statistical methods throughout your research project, from the planning phase to data analysis.

Contact: Sarah Kaspar (sarah.kaspar@embl.de)



Centre for Biomolecular Network Analysis (CBNA)

The CBNA disseminates expertise, know-how and guidance in network integration and analysis throughout EMBL.

Contact: Santhust (santhust@embl.de), Room 118

