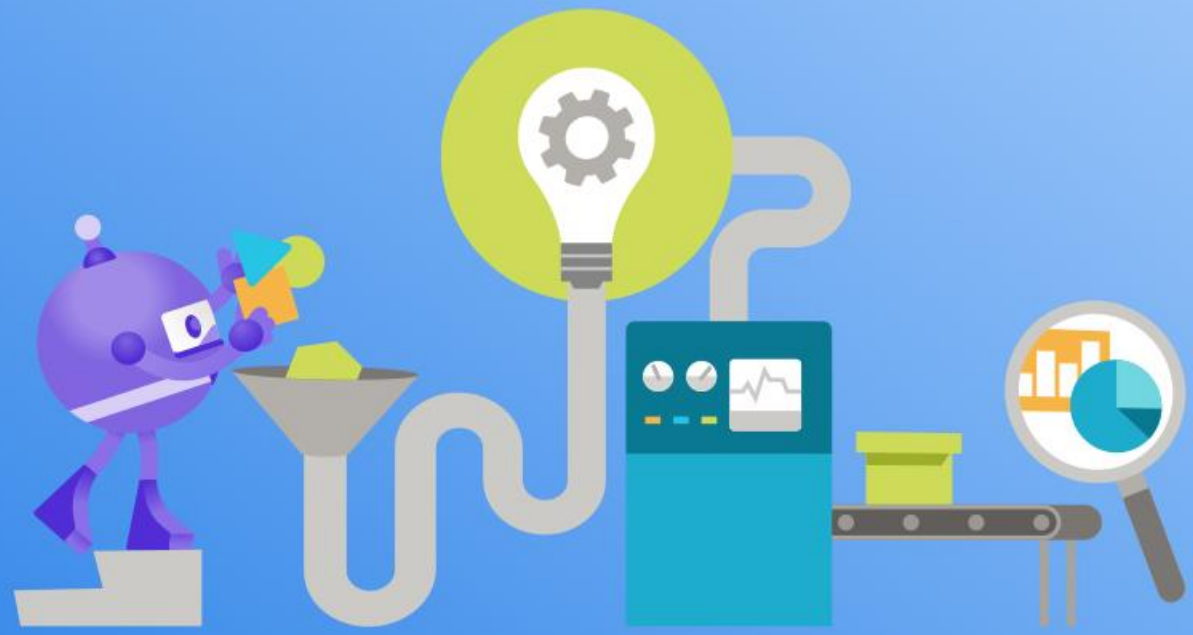


.NET Conf

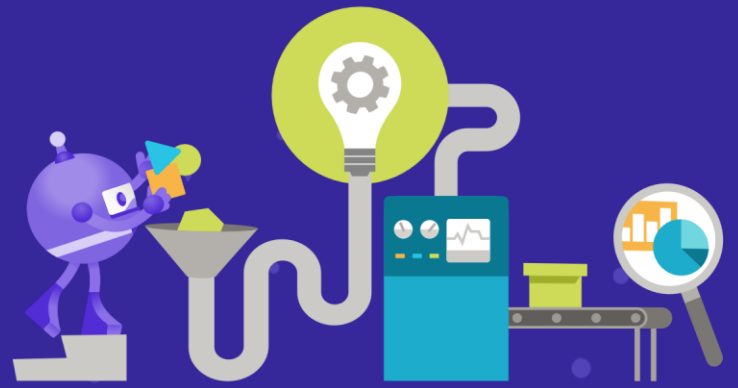
Focus on F#

July 29, 2021

focus.dotnetconf.net

Productive data science in biology with F#

Kevin Schneider

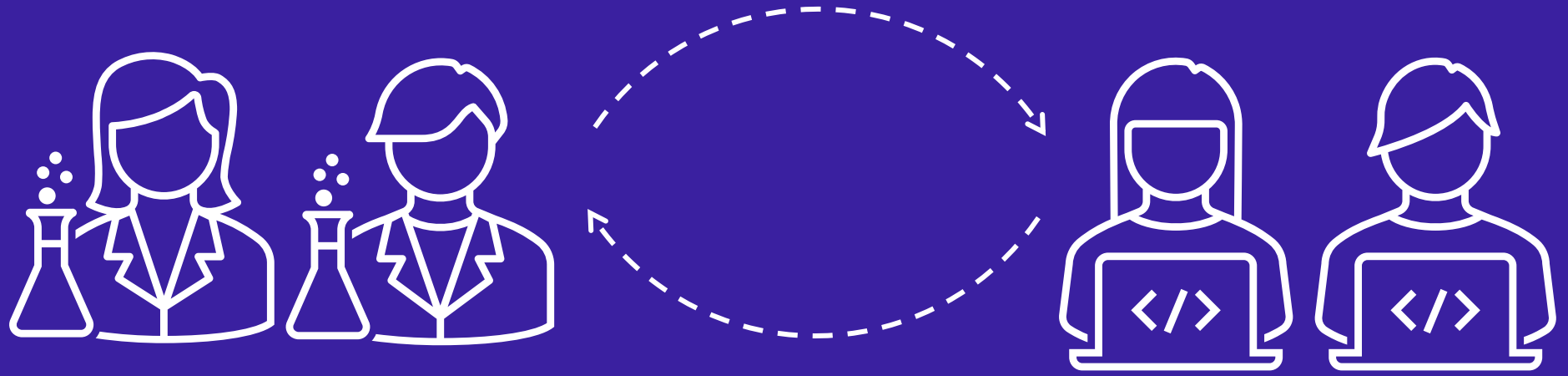


CSB COMPUTATIONAL SYSTEMS BIOLOGY

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

“Wet” lab

“Dry” lab

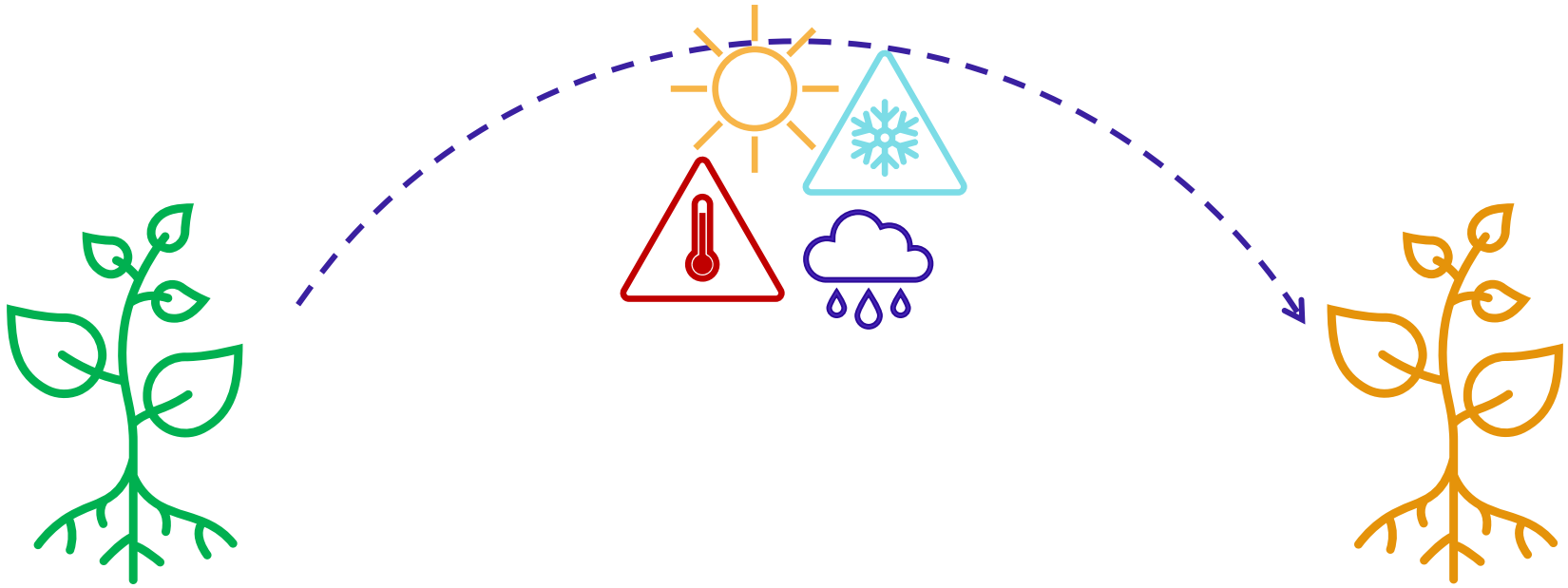


CSB

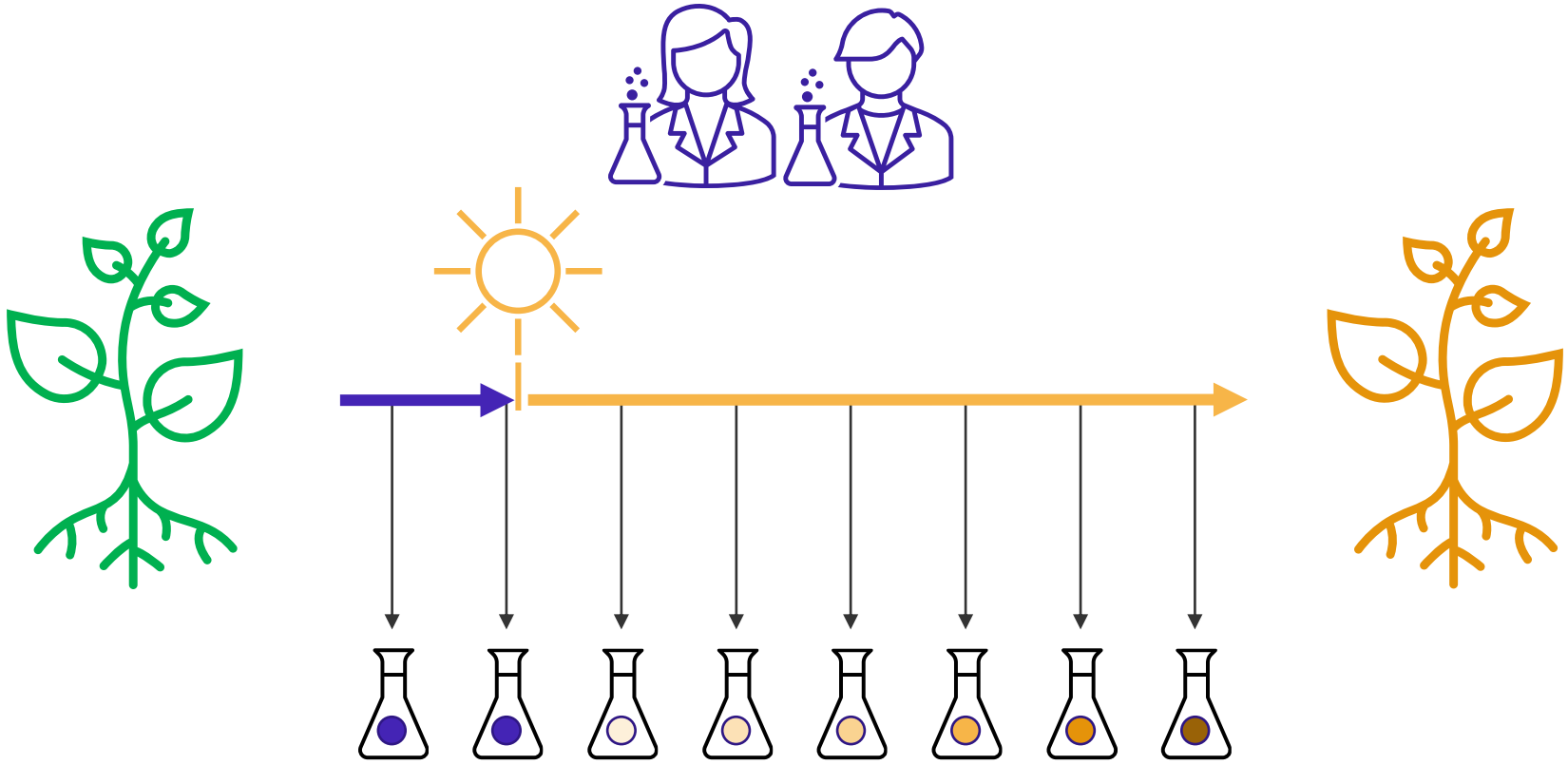
COMPUTATIONAL
SYSTEMS
BIOLOGY

TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

Plants need to react on a molecular level



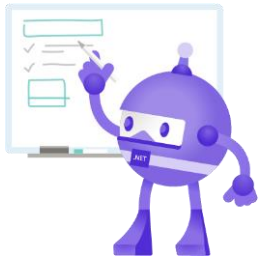
How to measure this response?



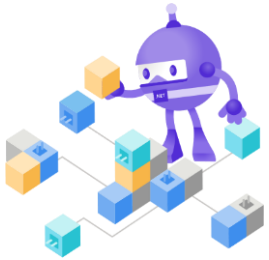
The pillars of our research activity

CSB COMPUTATIONAL SYSTEMS BIOLOGY

Teaching

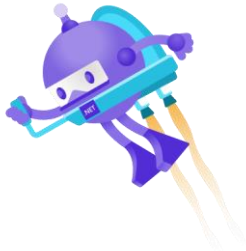


Method development

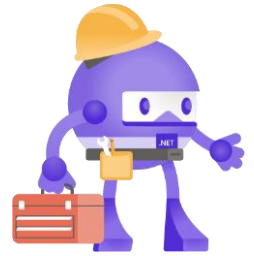


Data Analysis

Web Services



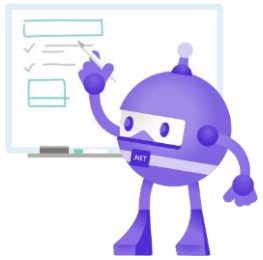
Open Science Tooling



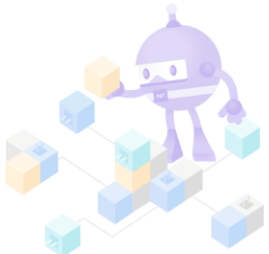
The pillars of our research activity

CSB COMPUTATIONAL SYSTEMS BIOLOGY

Teaching



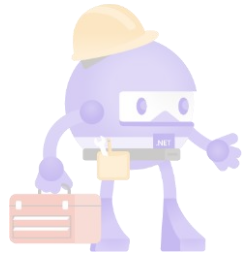
Method development



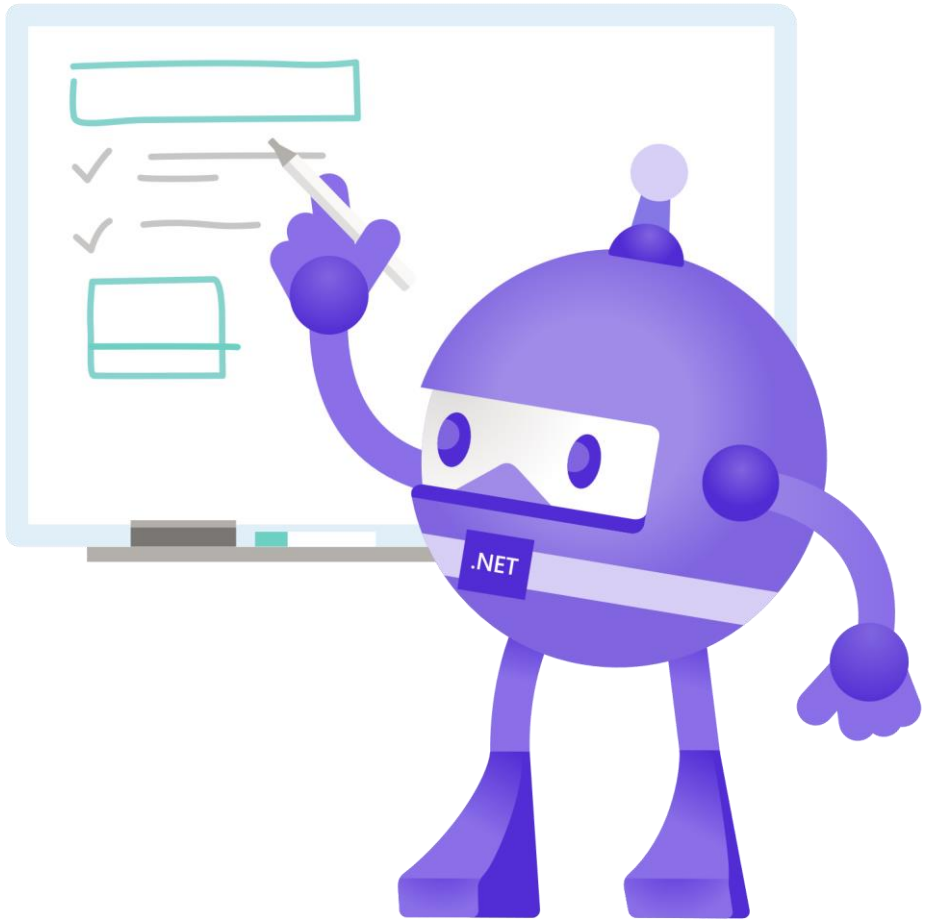
Data Analysis

Web Services

Open Science Tooling



Teaching



- F# from the ground up for scientific programming
- data analysis and project work

⇒ Robust tooling:
  - fsdocs (FSharp.Formatting)
  - dotnet interactive notebooks

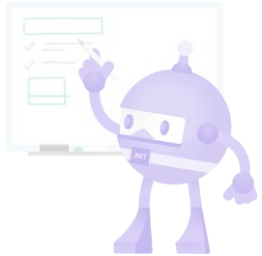
Demo

fsdocs and dotnet interactive notebooks for digital teaching

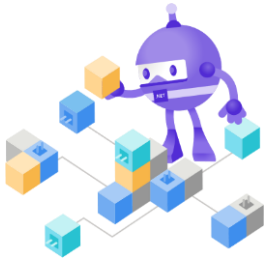
The pillars of our research activity

CSB COMPUTATIONAL SYSTEMS BIOLOGY

Teaching

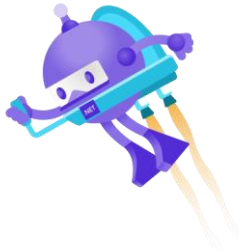


Method development

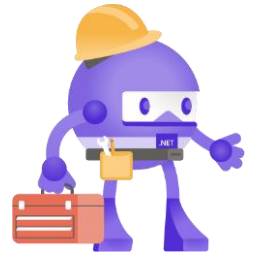


Data Analysis

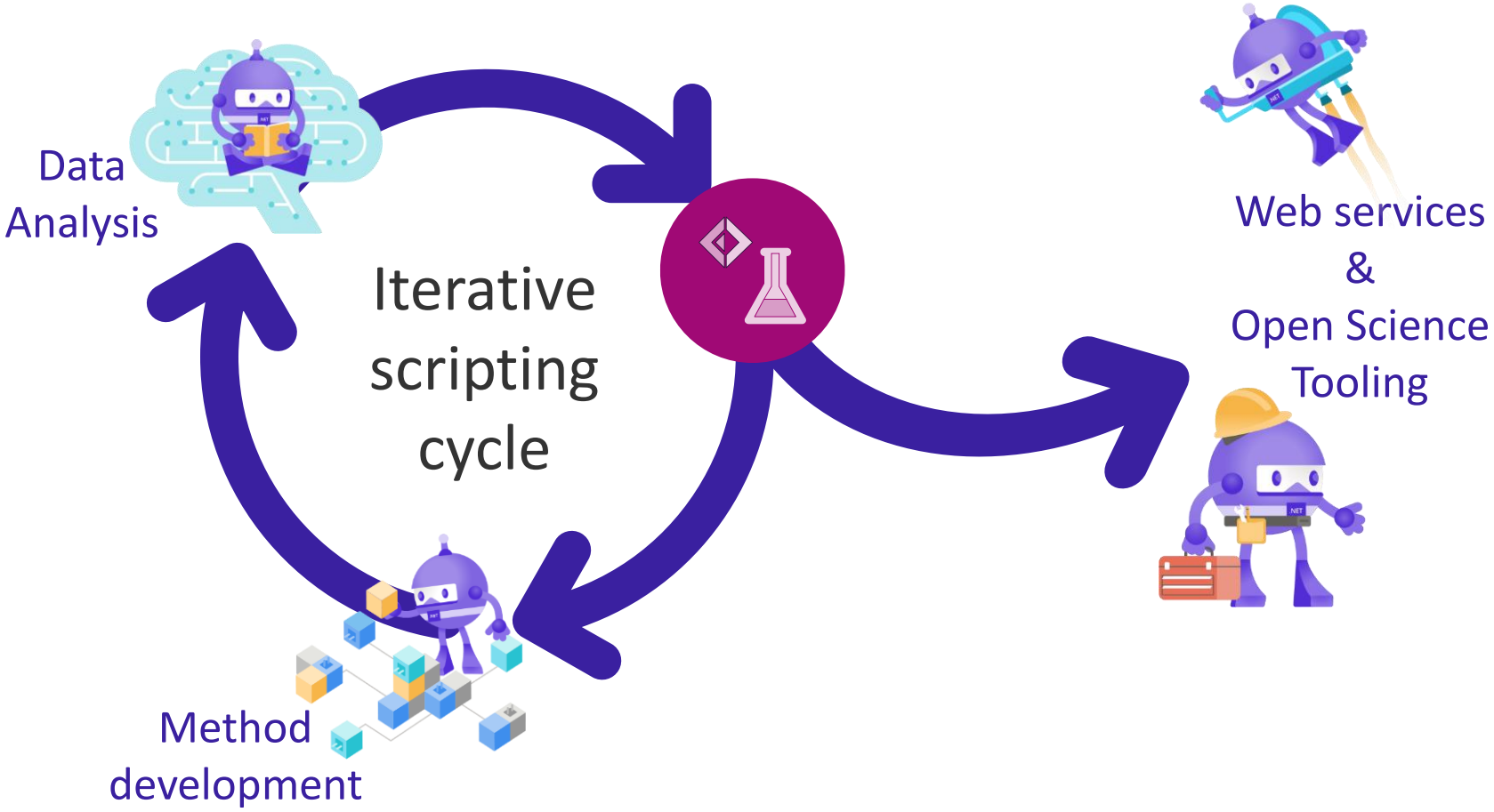
Web Services



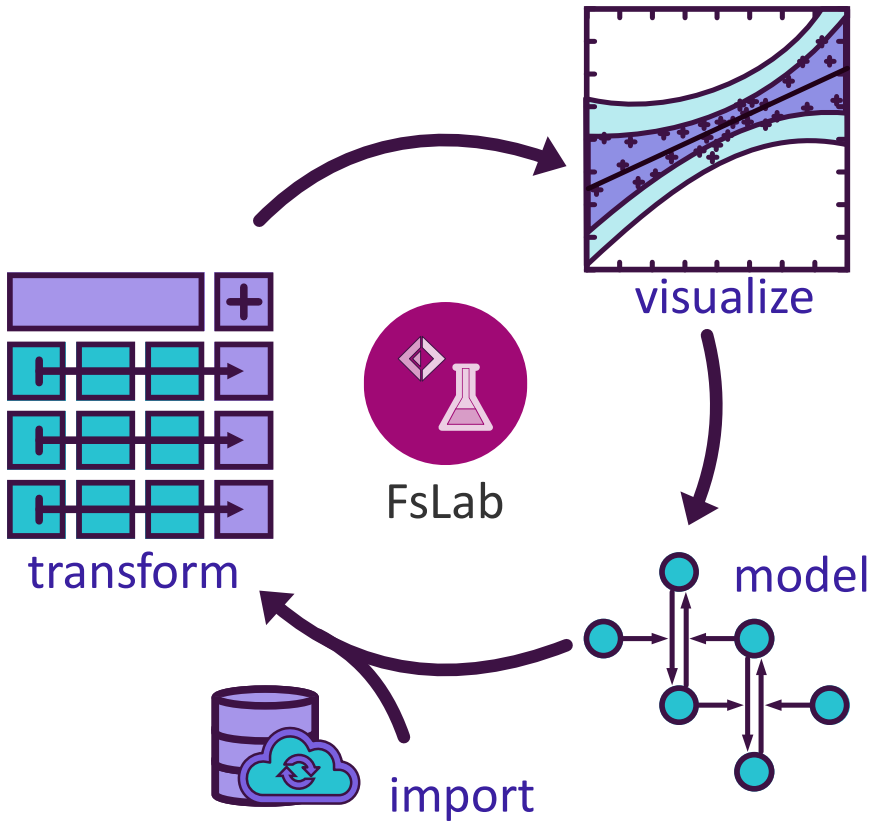
Open Science Tooling



From method development to tooling for open science



FsLab: Perform the whole data science cycle in F#!



Deedle

Plotly.NET

Cyjs.NET

FSharp.Stats

flips

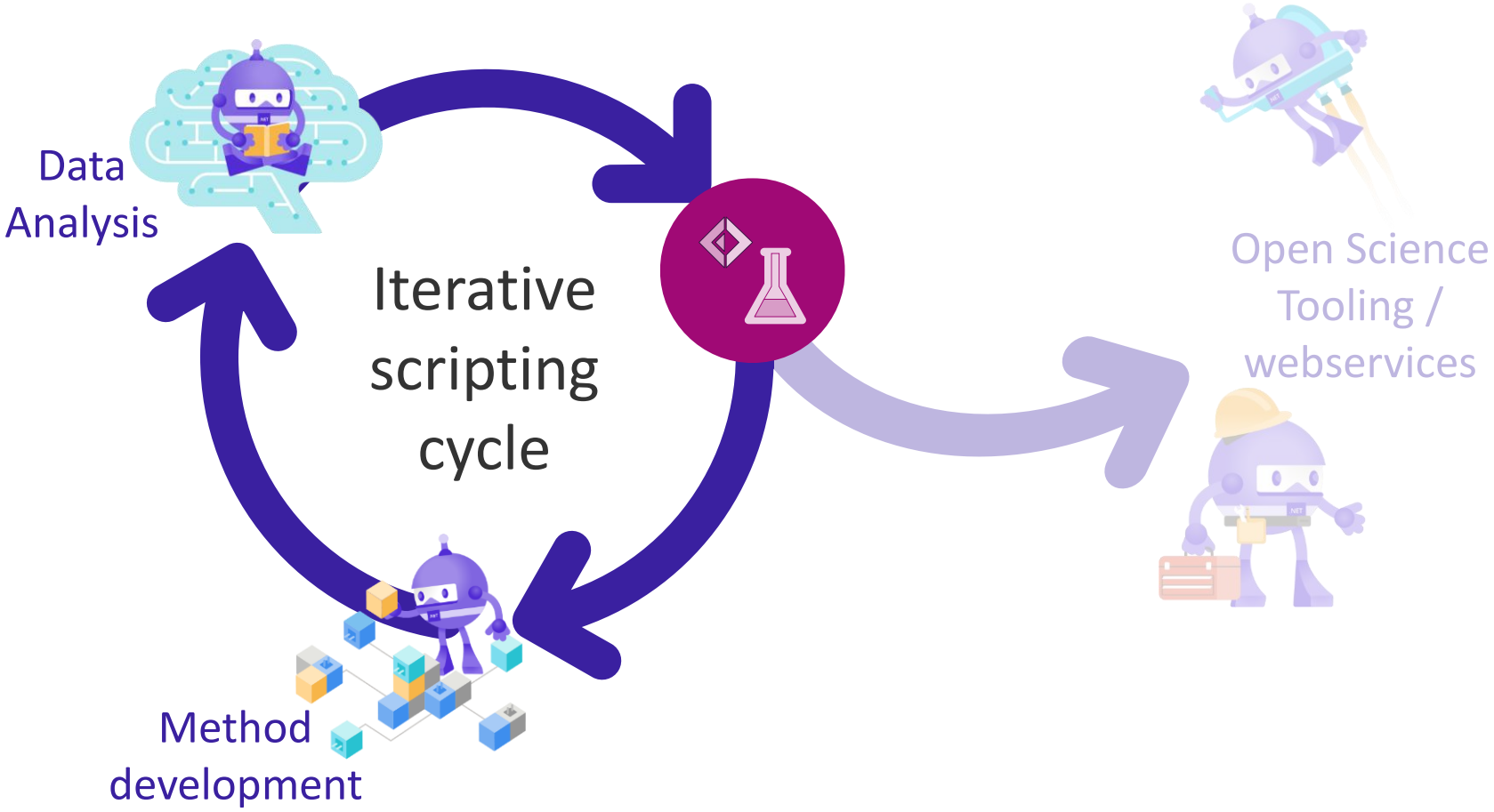
FSharp.Data

ML.NET

SciSharp STACK

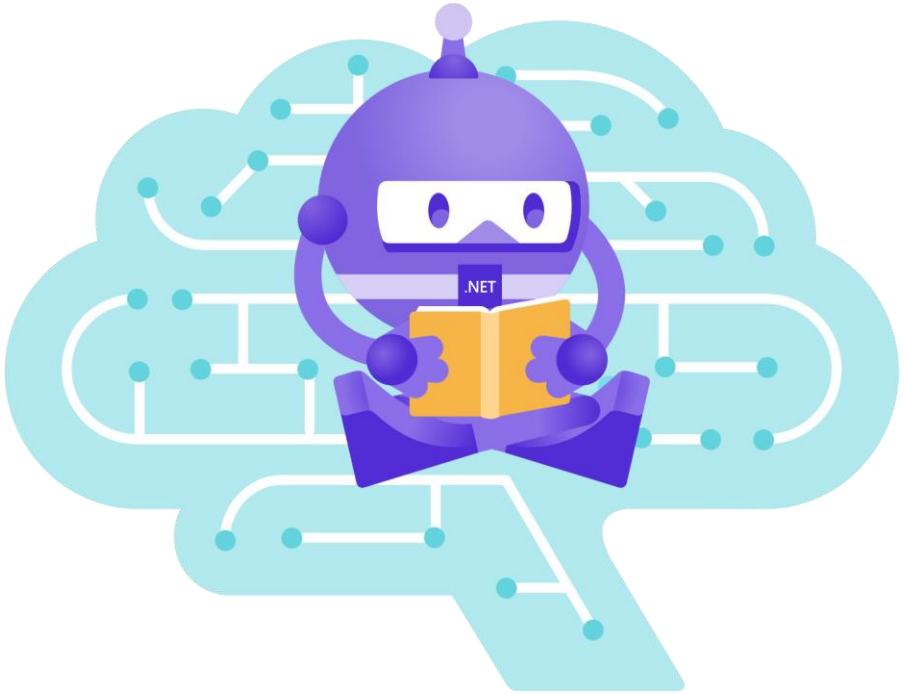
https://fslab.org

From method development to tooling for open science

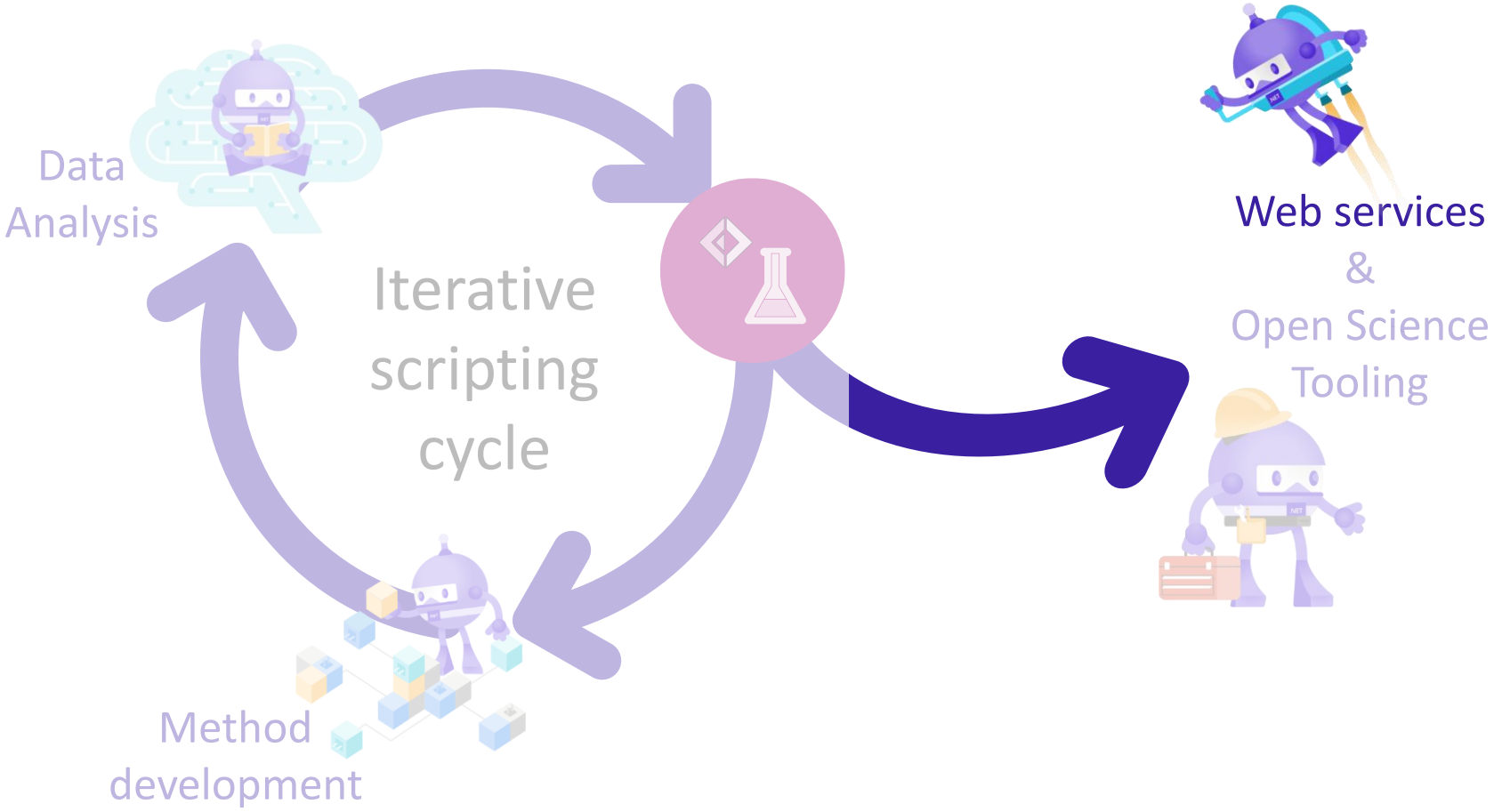


Demo

FsLab and data analysis pipeline flow



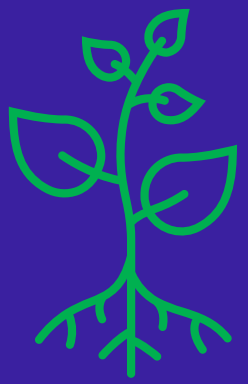
From method development to tooling for open science



Web Services

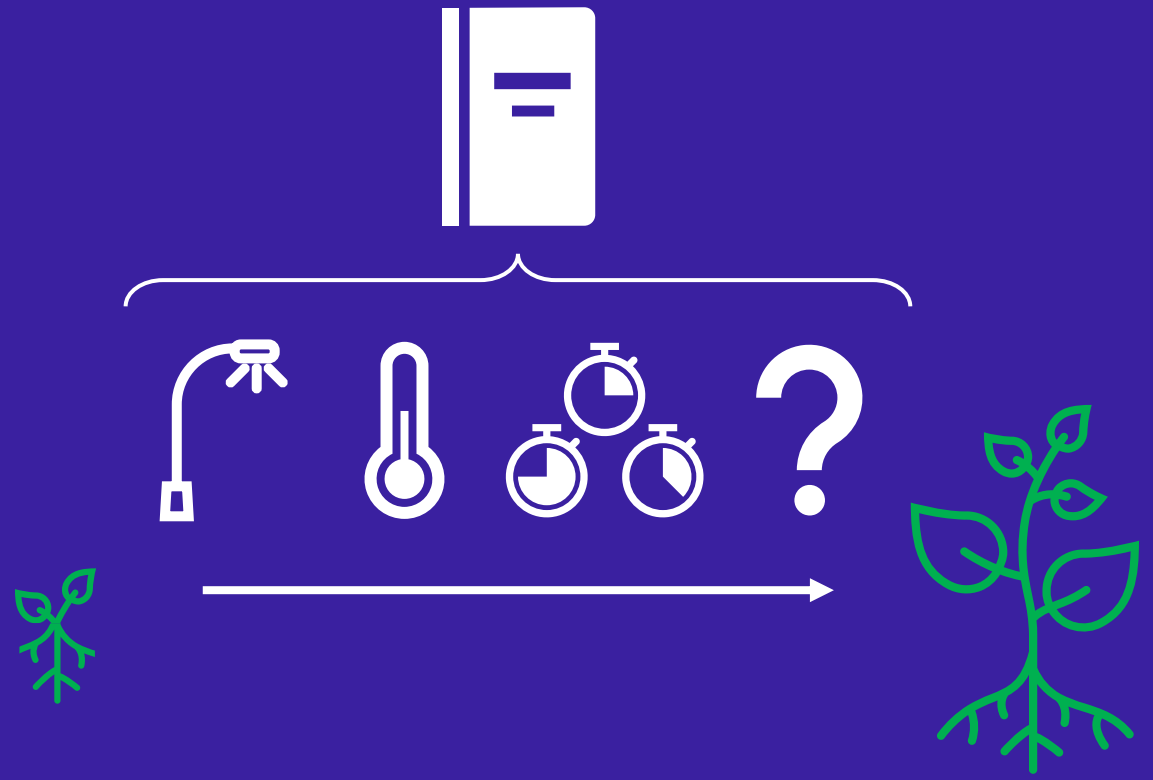
- SAFE Stack / Fable
  ⇒ mature ecosystem
- Data Analysis pipelines:
  - You provide input
  - We run the pipeline and return the result
- Office add-ins

“Wet” lab



“Wet” lab

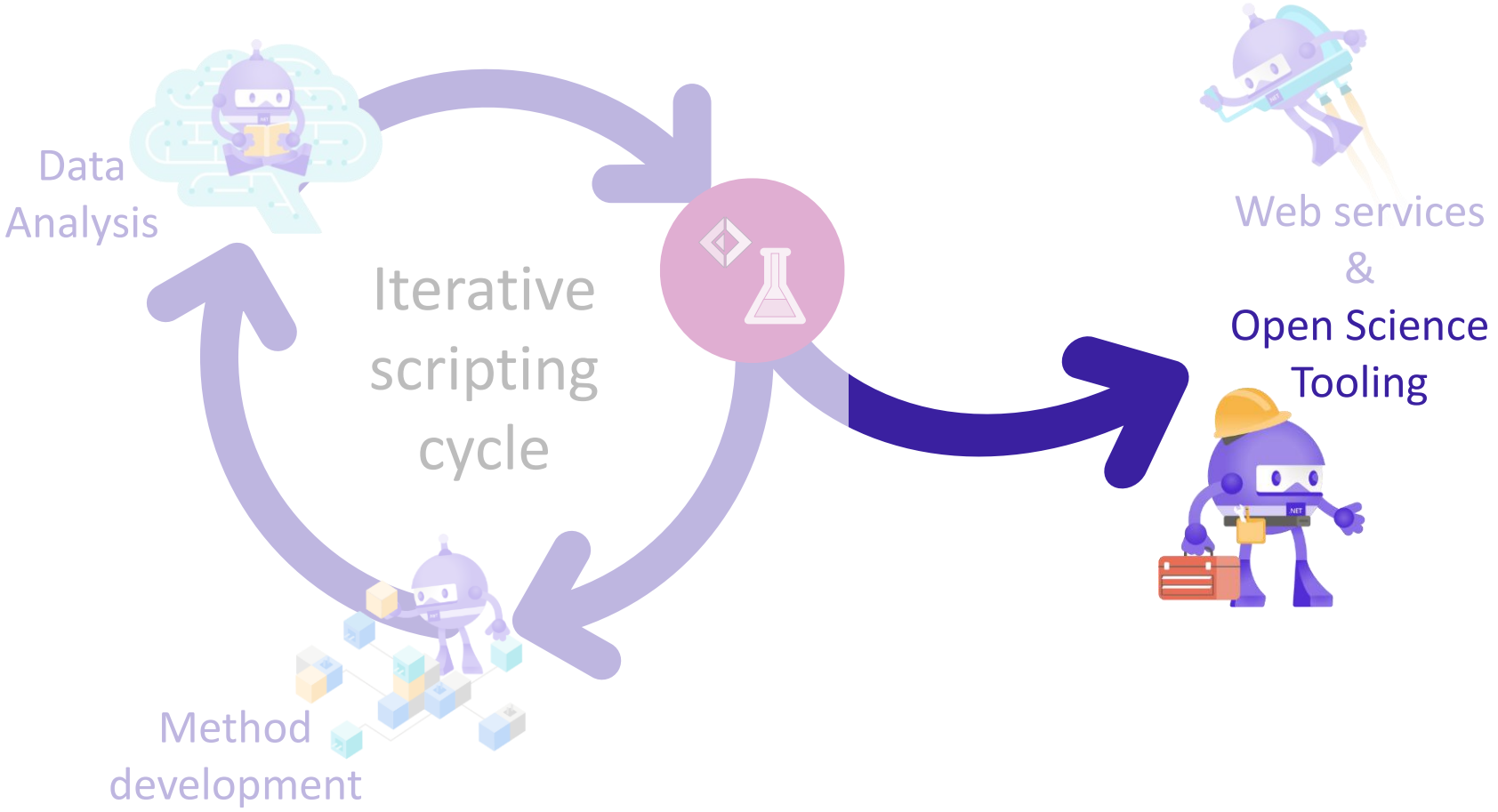
“Wet” lab



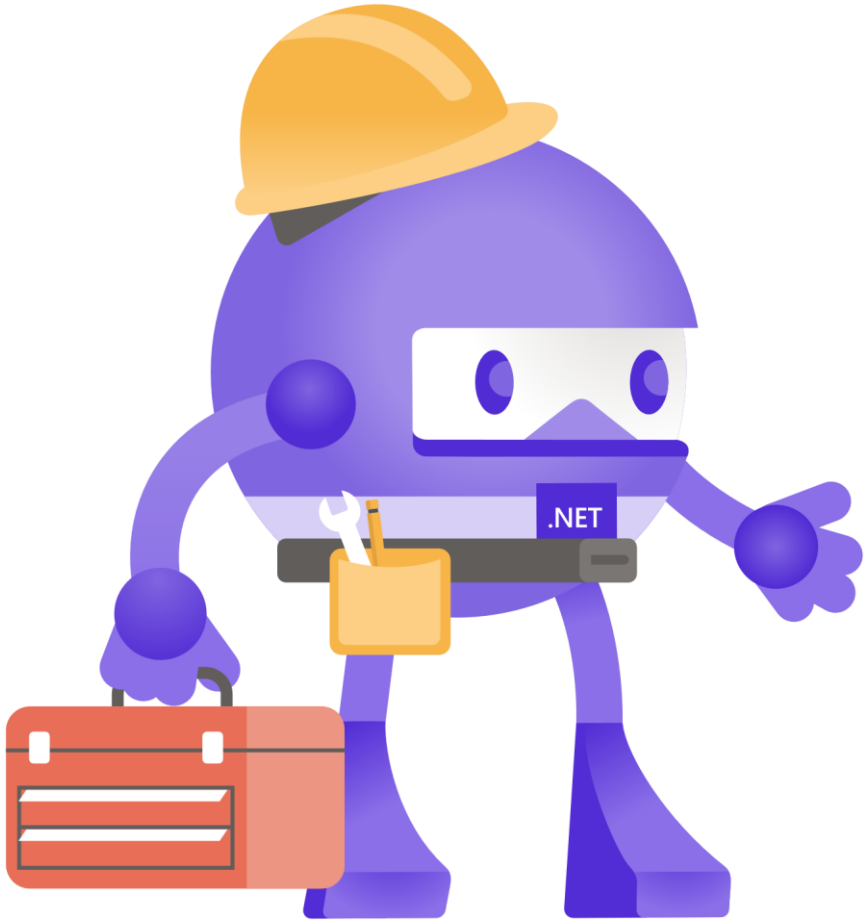
Demo

SWATE excel add-in

From method development to tooling for open science



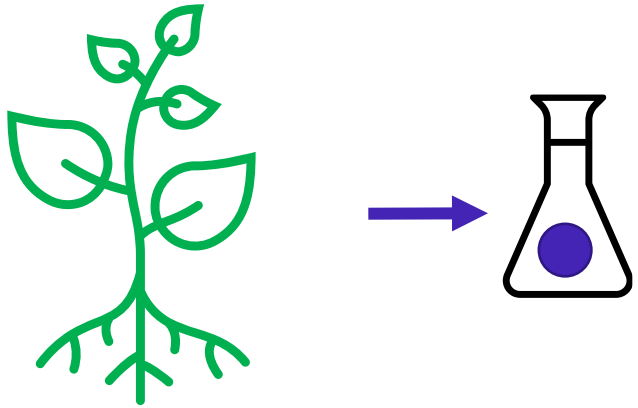
Open Science Tooling



- Command line tools
- Docker containers
- Enabling everyone to perform complex data analysis tasks:

=> dotnet on usegalaxy.eu

Proteomics



Proteomics

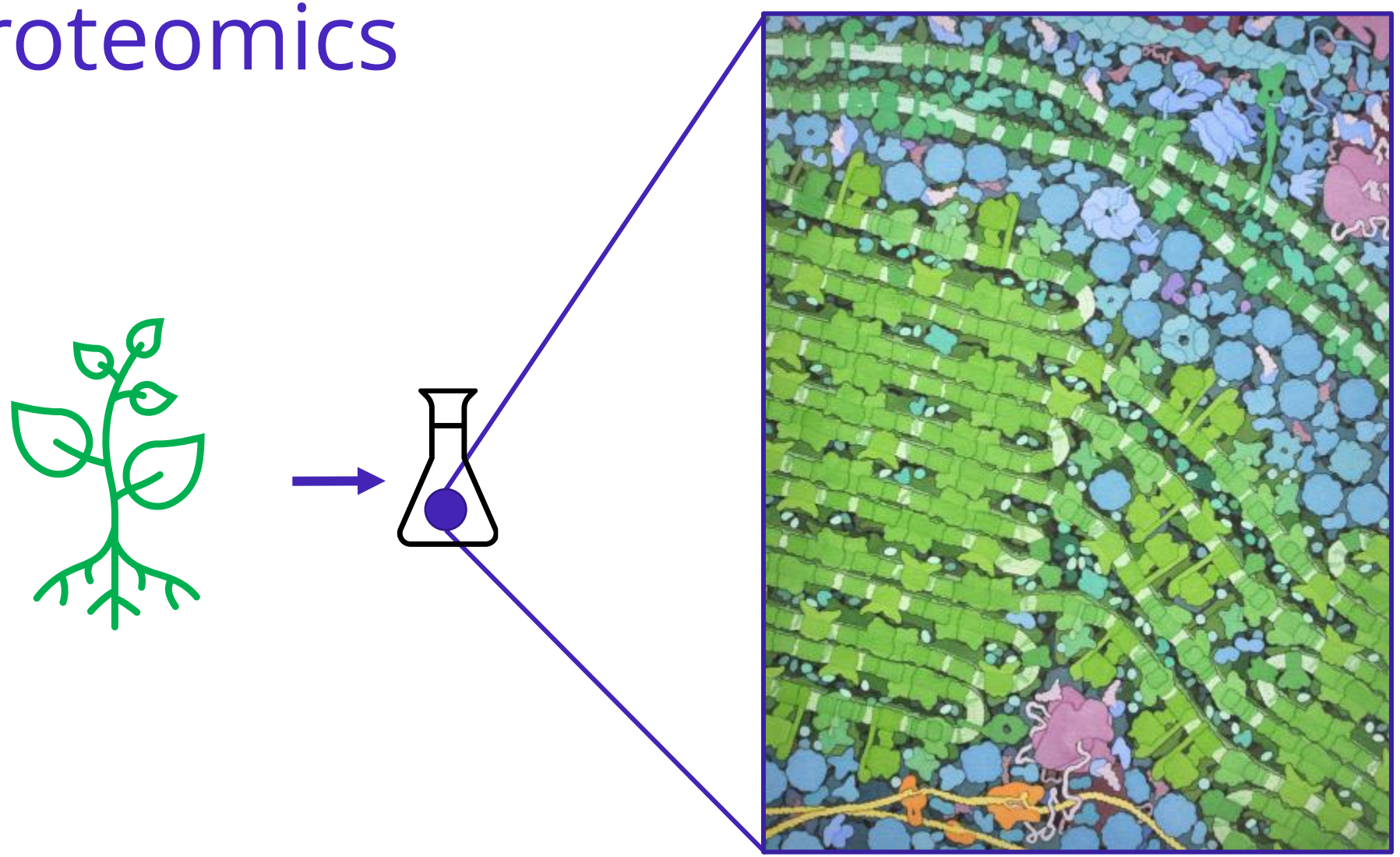
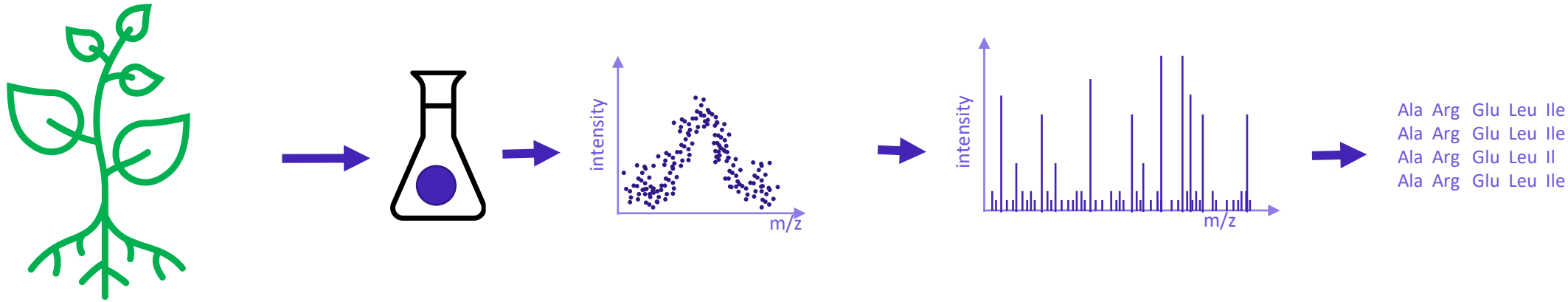
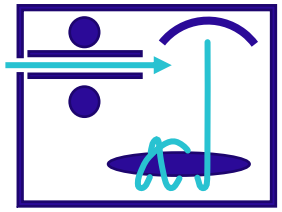


Illustration by David S. Goodsell.
doi: 10.2210/rcsb_pdb/goodsell-gallery-010

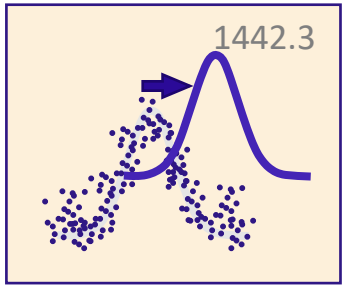
Proteomics


intensity
m/z
intensity
m/z
Ala Arg Glu Leu Ile
Ala Arg Glu Leu Ile
Ala Arg Glu Leu Il
Ala Arg Glu Leu Ile

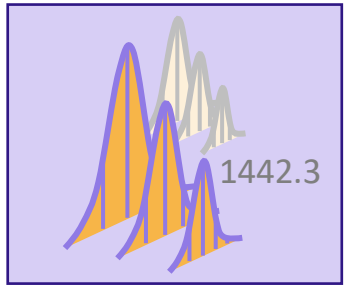
Proteomics

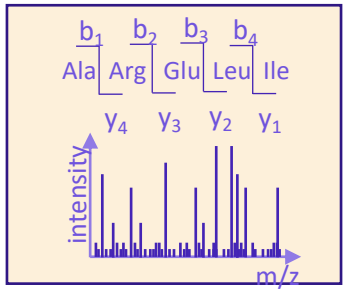


signal detection

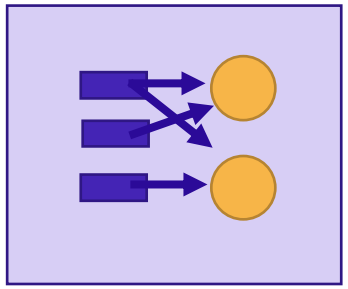


quantification

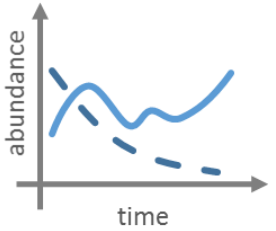




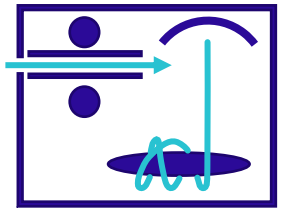
identification



protein inference



Proteomics



signal detection

quantification

ProteomIQon:

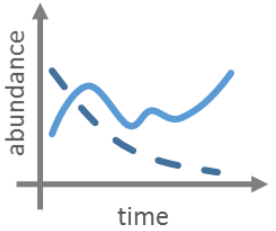
Full Proteomics pipeline

written in F#

packaged as dotnet CLI tools

identification

protein inference



Proteomics

- Computational pipeline is complex and consist of many steps
  => many different tools and parameters

- Users across all fields of biology and adjacent sciences

=> We need a HPC capable platform with a user-friendly GUI for everyone

Galaxy PROJECT

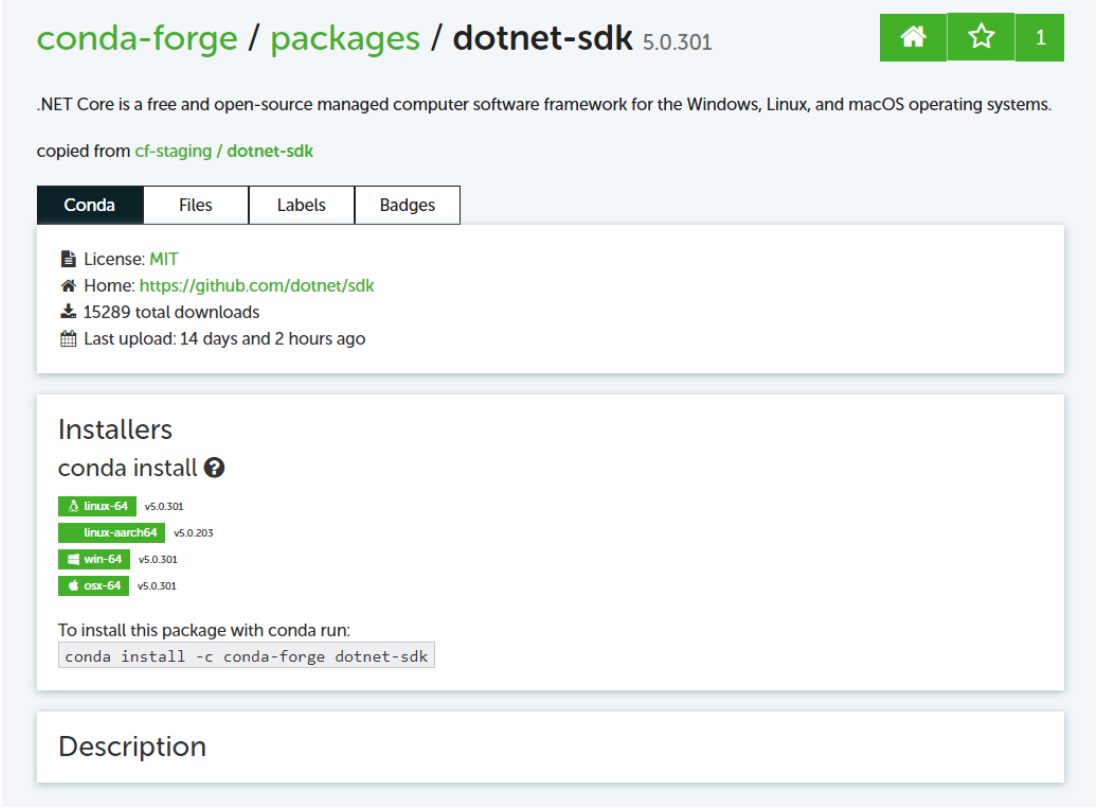
The galaxy project

- open, web-based platform
- data-intensive computational research

- Run and repeat analysis workflows
- Share results

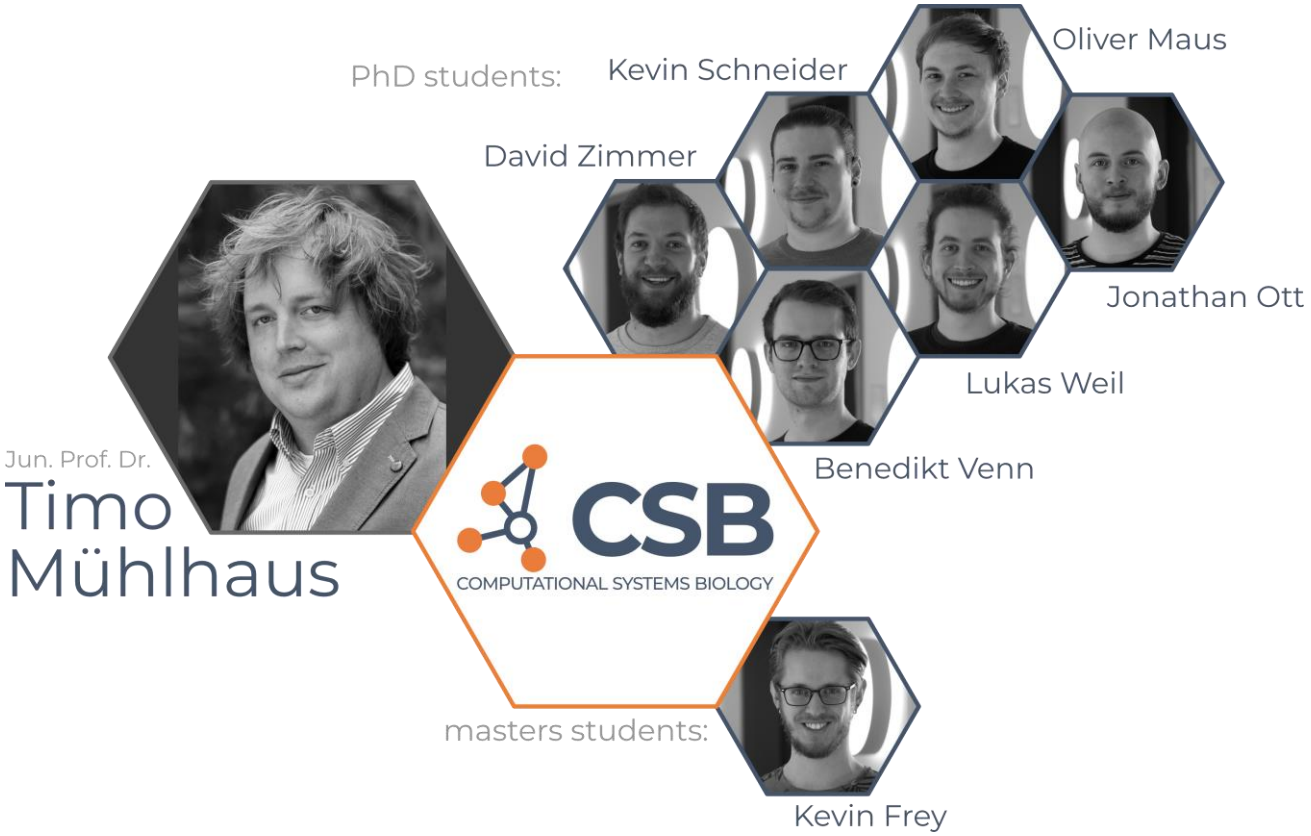
FSharp on galaxy

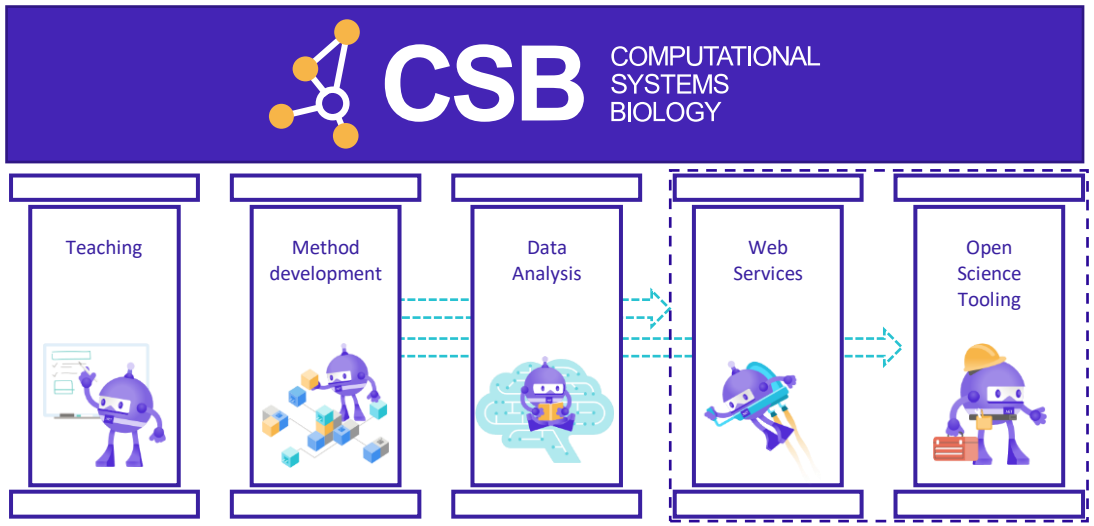
- Together we brought dotnet to conda
- You can now package dotnet tools as conda packages
- These packages can be used on the galaxy platform



Demo

Dotnet tools on galaxy


PhD students:
Kevin Schneider
Oliver Maus
David Zimmer
Jonathan Ott
Lukas Weil
Benedikt Venn
Jun. Prof. Dr.
Timo
Mühlhaus
CSB
COMPUTATIONAL SYSTEMS BIOLOGY
masters students:
Kevin Frey


CSB
COMPUTATIONAL
SYSTEMS
BIOLOGY
Teaching
Method
development
Data
Analysis
Web
Services
Open
Science
Tooling

Find me

@kMutagene
@kMutagene

Find us

@CSBiology
@cs_biology
csb.bio.uni-kl.de

@fslaborg
@fslaborg
fslab.org

Thanks for joining!

Ask your questions live on Twitter #dotNETConf