

Tool deployment and sustainability in bioinformatics

✓ Fixed

Bérénice Batut, Jonas Weber, Rolf Backofen, Björn Grüning

University of Freiburg

Bioinformatics for Human Health and Disease
Heidelberg - November 2016

Biomedical research

Generation of massive amounts of diverse data

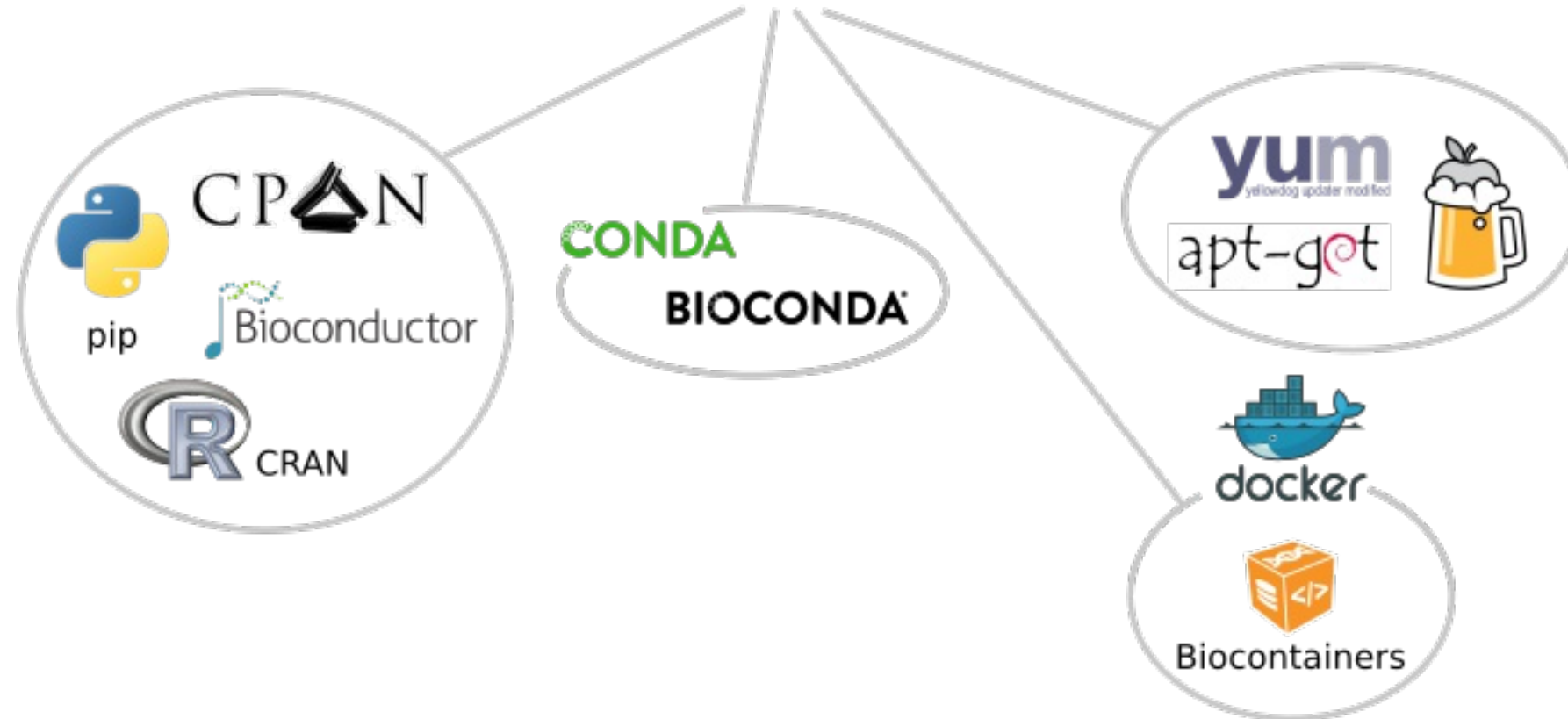
→ Need of bioinformatic solutions and tools

How to develop a tool?

Question → Development of code → Deployment

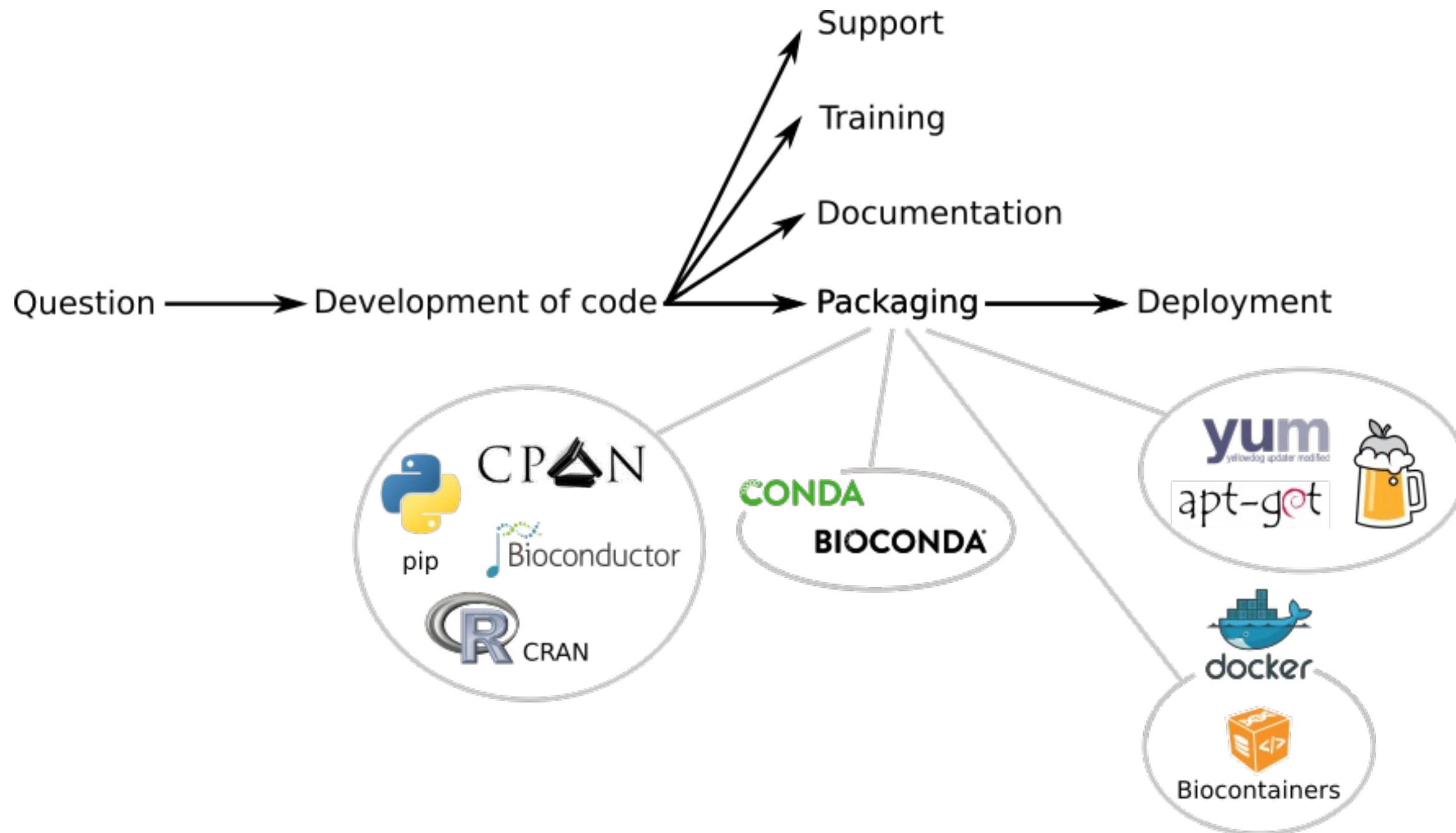
How to develop a tool?

Question → Development of code → Packaging → Deployment




How to develop a tool?

The Golden Path



But...

Who has never dealt with missing tool dependencies?
Or an older version of a tool that could not be installed?

 Major issues in
deployment and sustainability

Deployment issue

Wanted

A package manager

- Programming language agnostic
- OS independent
- No root privileges needed
- Management of multiple version
- HPC and Cloud compatible

A great solution



- Open Source package manager
- Independent of any programming language and OS
- Fast, robust and easy package installation

```
$ conda install deeptools
```

- Installation and management of multiple versions

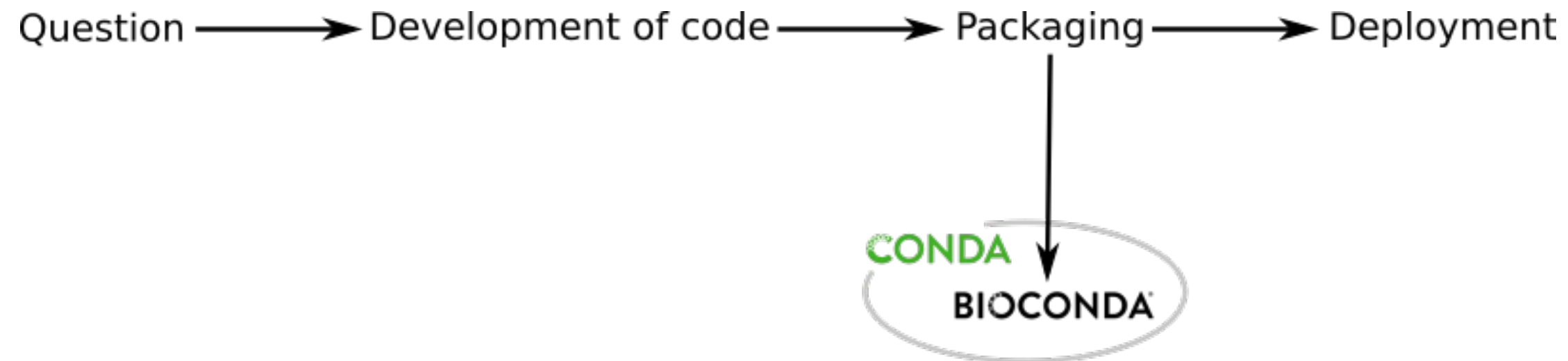
A Conda package



→ Easy to write!

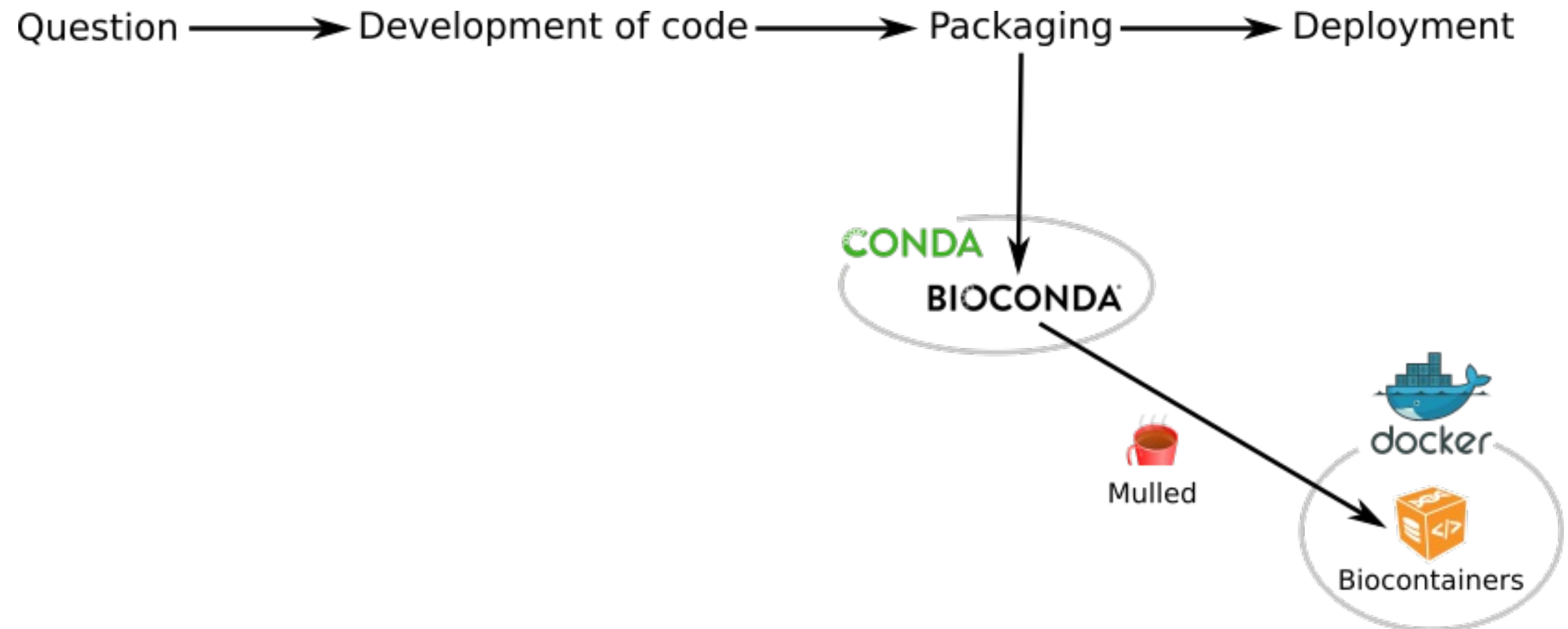
Bioinformatics tools

BIOCONDA[®]



- More than 1,700 bioinformatic packages
- Big, fast-growing and newcomer-friendly community

Want to have a Docker image?

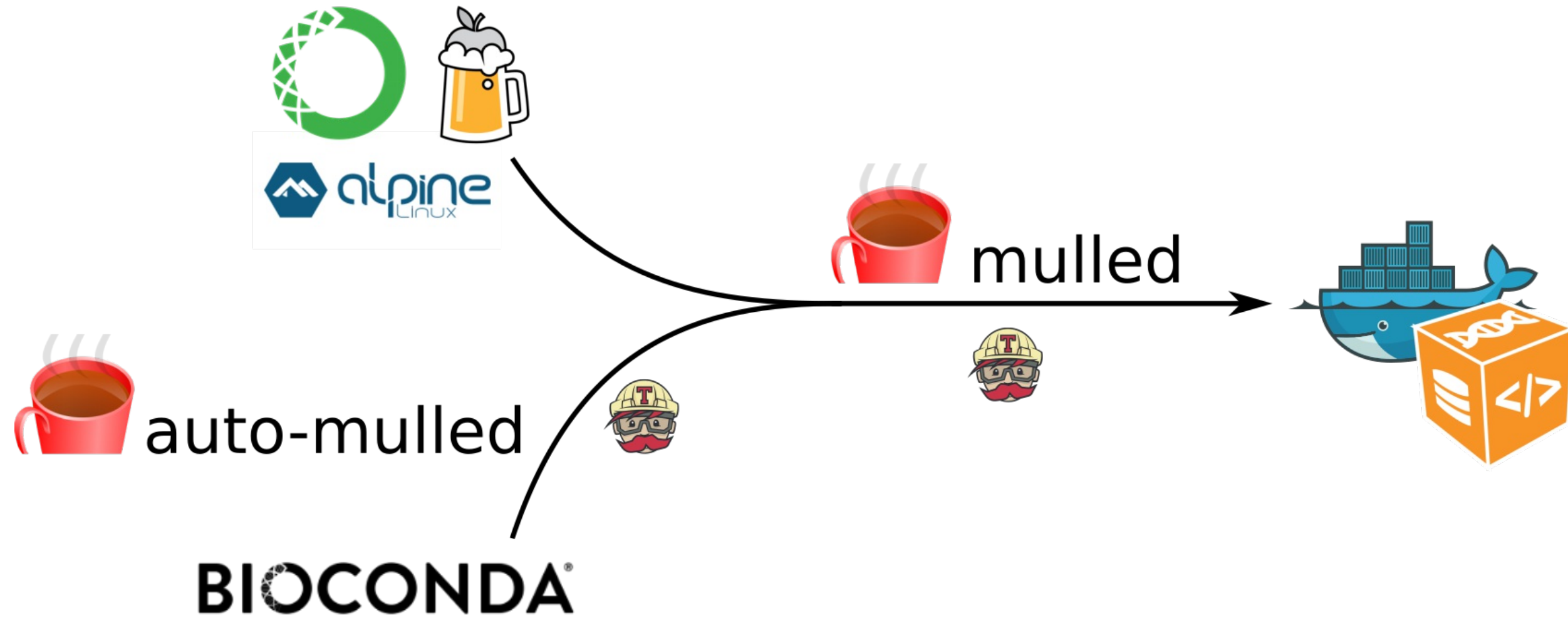


Mulled

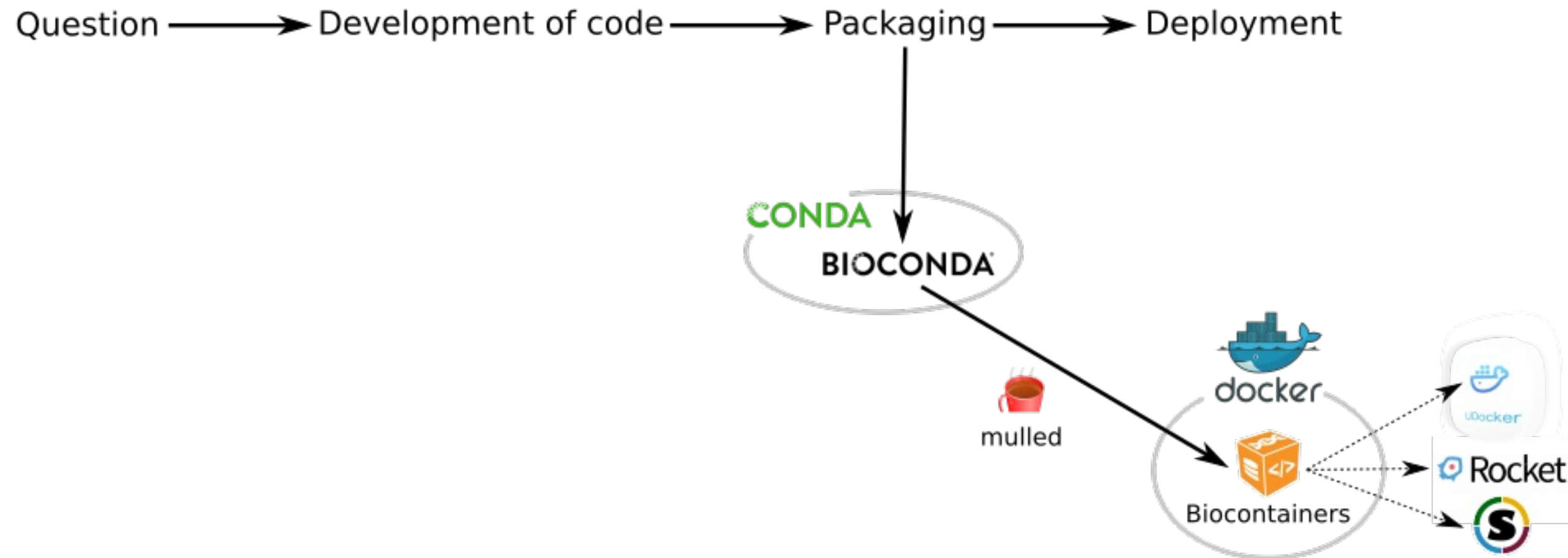
A layer donning approach to build Docker containers
(without Dockerfile)



Mulled



Deployment issue → ✓ Fixed!



Sustainability issue

Wanted

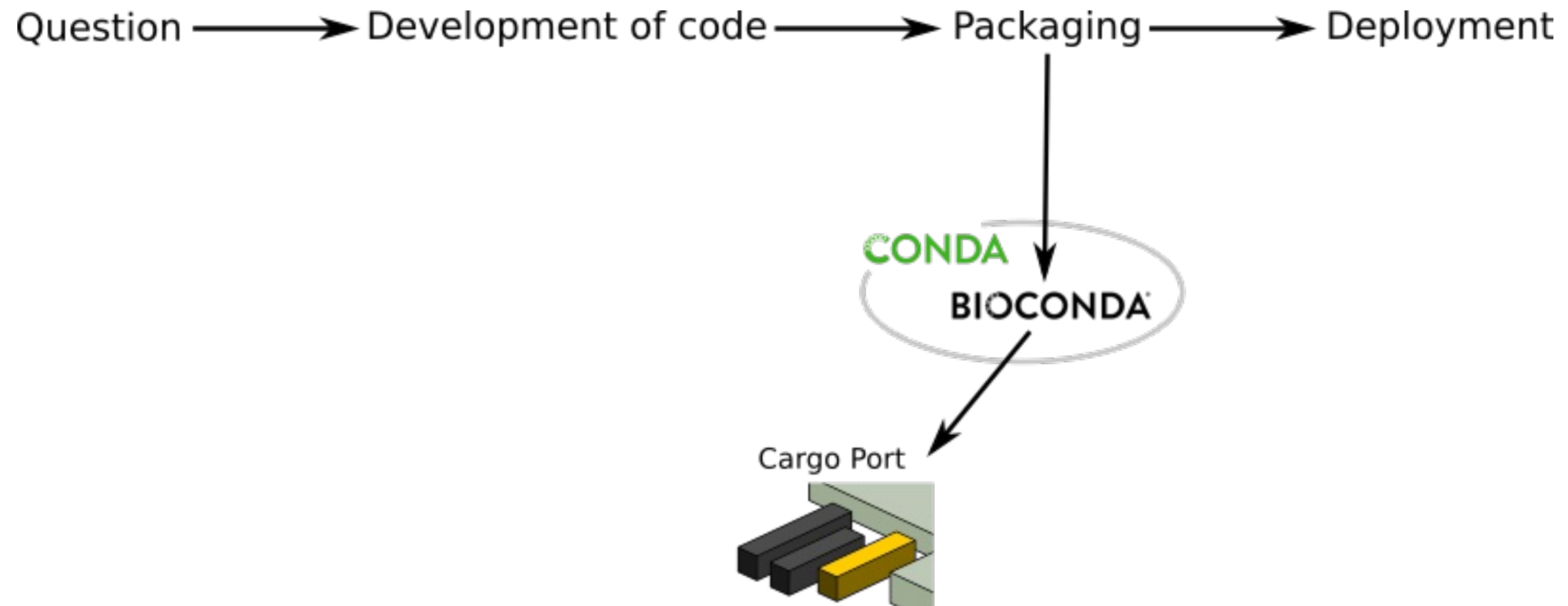
- Accessibility of all packages
- Cache of all versions of the packages
- Permanent access

A solution

Cargo Port

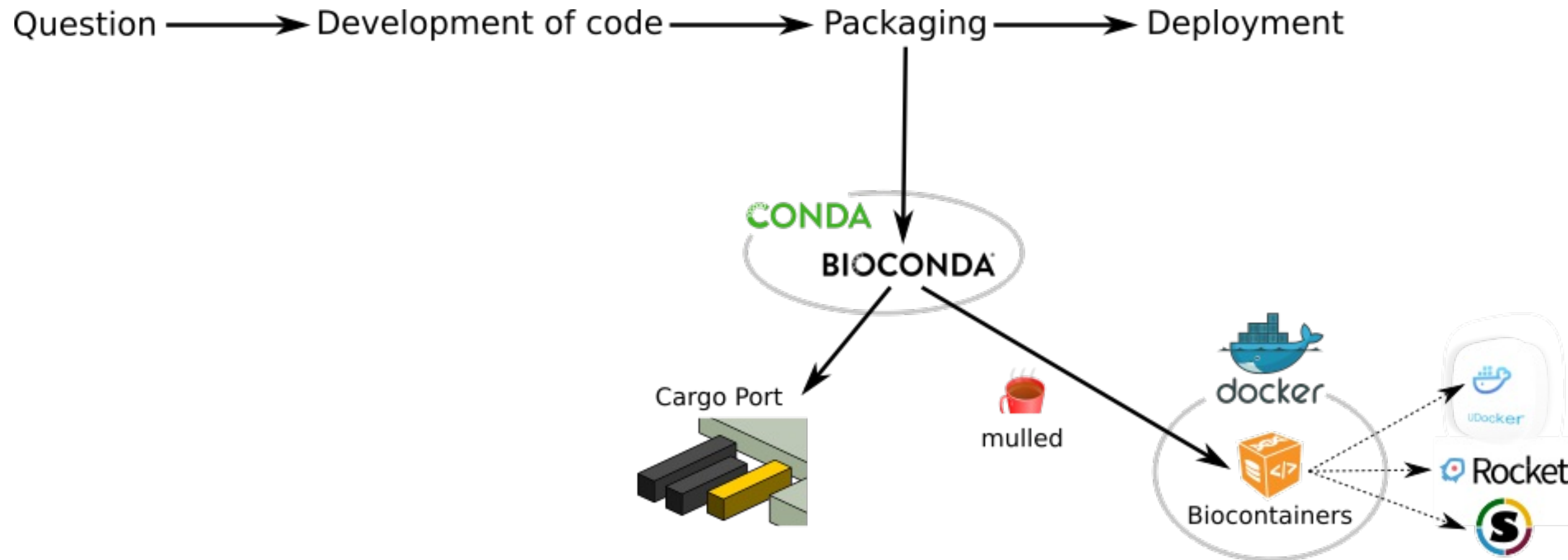
- [A public repository](#), kindly hosted by the Galaxy Project
- Stable and long term storage
- Daily archives of the BioConda packages

Sustainability issue → ✓ Fixed!



Tool deployment and sustainability in bioinformatics

→ ✓ Fixed!



Thank You!

Questions?

