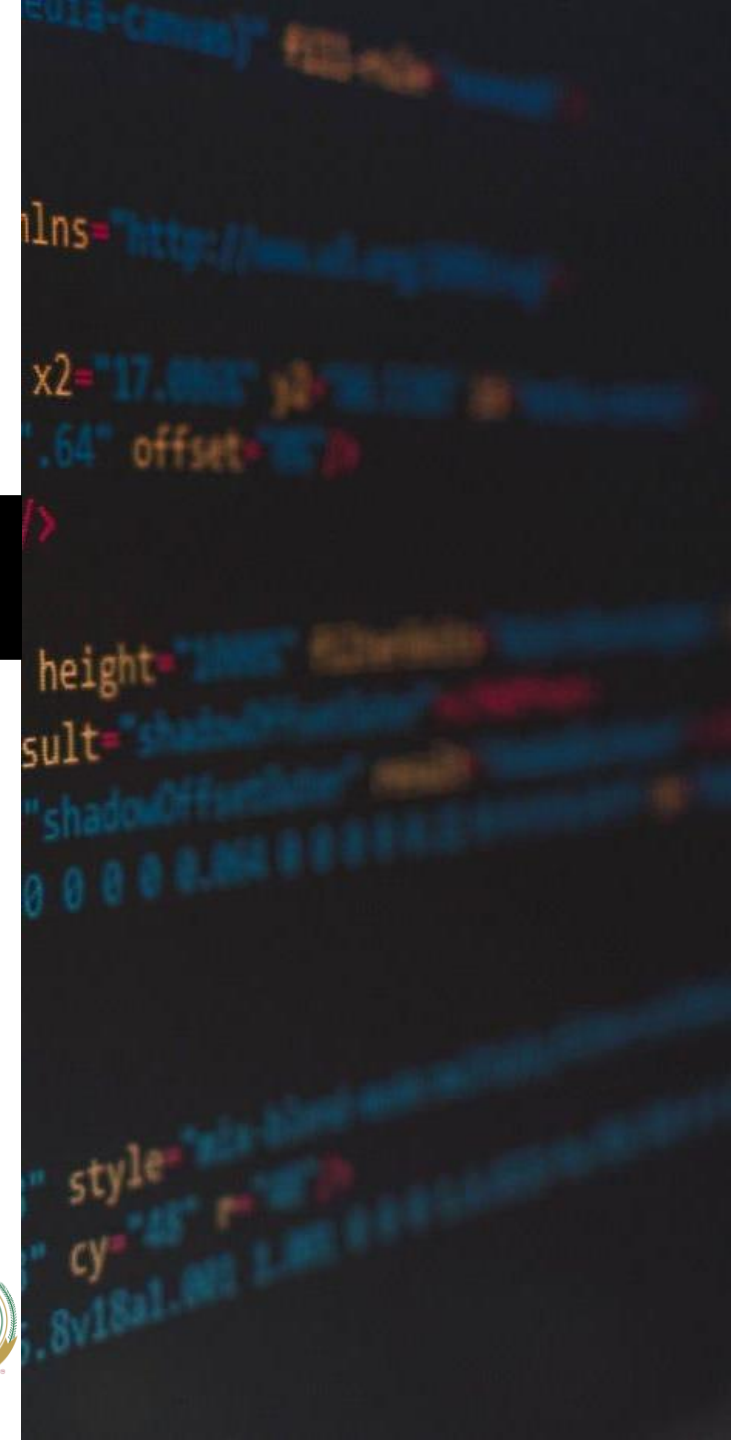


# INTRODUCTION TO CONDA & ENVIRONMENTS

*Klebsiella Workshop*

*Sep 2024*



SANBI

South African National  
Bioinformatics Institute



UNIVERSITY of the  
WESTERN CAPE



**ASLM**  
AFRICAN SOCIETY FOR LABORATORY MEDICINE

AFRICA CDC  
Centres for Disease Control and Prevention  
Safeguarding Africa's Health



PUBLIC HEALTH ALLIANCE FOR  
GENOMIC EPIDEMIOLOGY

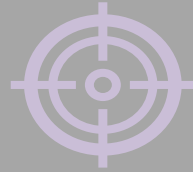
African  
Union



# Contents



Introduction



Aims



Conda & Environments



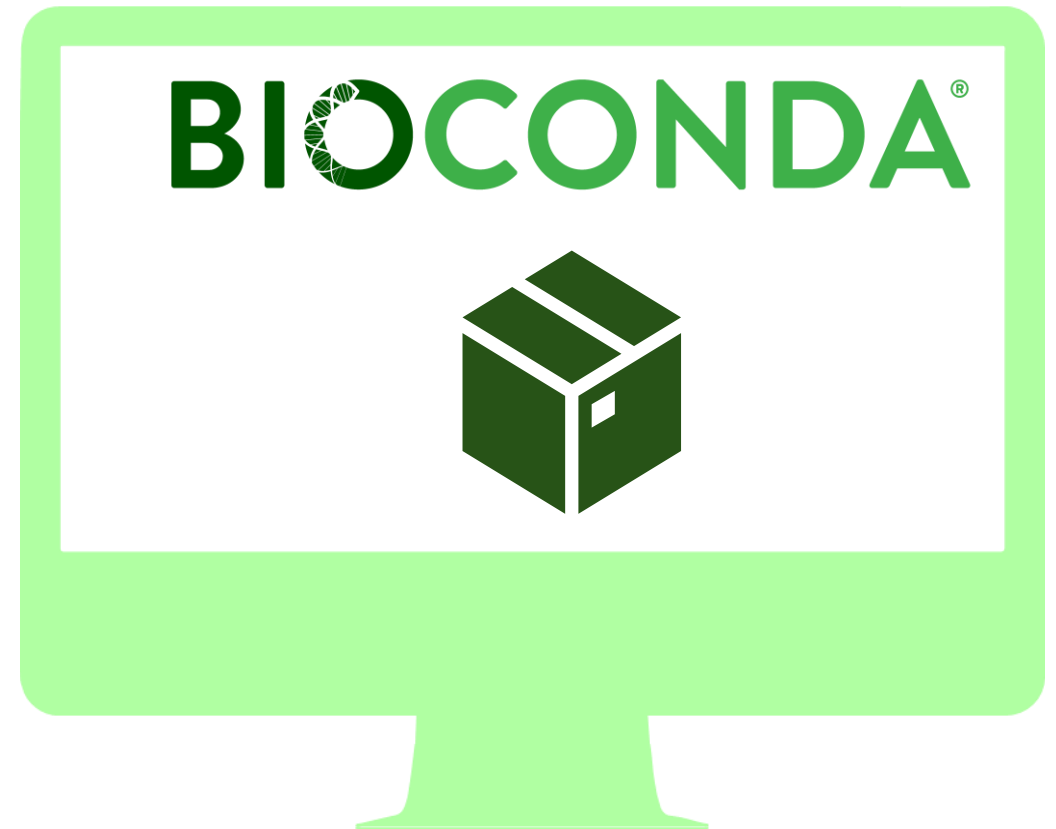
Getting Started



Installing & Running Kleborate

# Aims

- Introduce the advantages of conda & bioinformatics
- Basics of conda commands
- Implementation of commands
- Installation & execution of tools
- Complete hands-on exercise



# Introduction

YOU ARE HERE

INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS



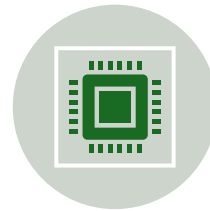
CONDA IS A  
PACKAGE  
MANAGER  
APPLICATION  
RUN ON THE  
COMMAND  
LINE



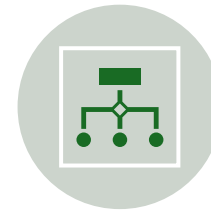
ALLOWS FOR  
QUICKLY  
INSTALLING,  
RUNNING AND  
UPDATING  
PACKAGES



TOOLS AND  
PIPELINES CAN  
BE EASILY  
DEVELOPED IN  
CONDA



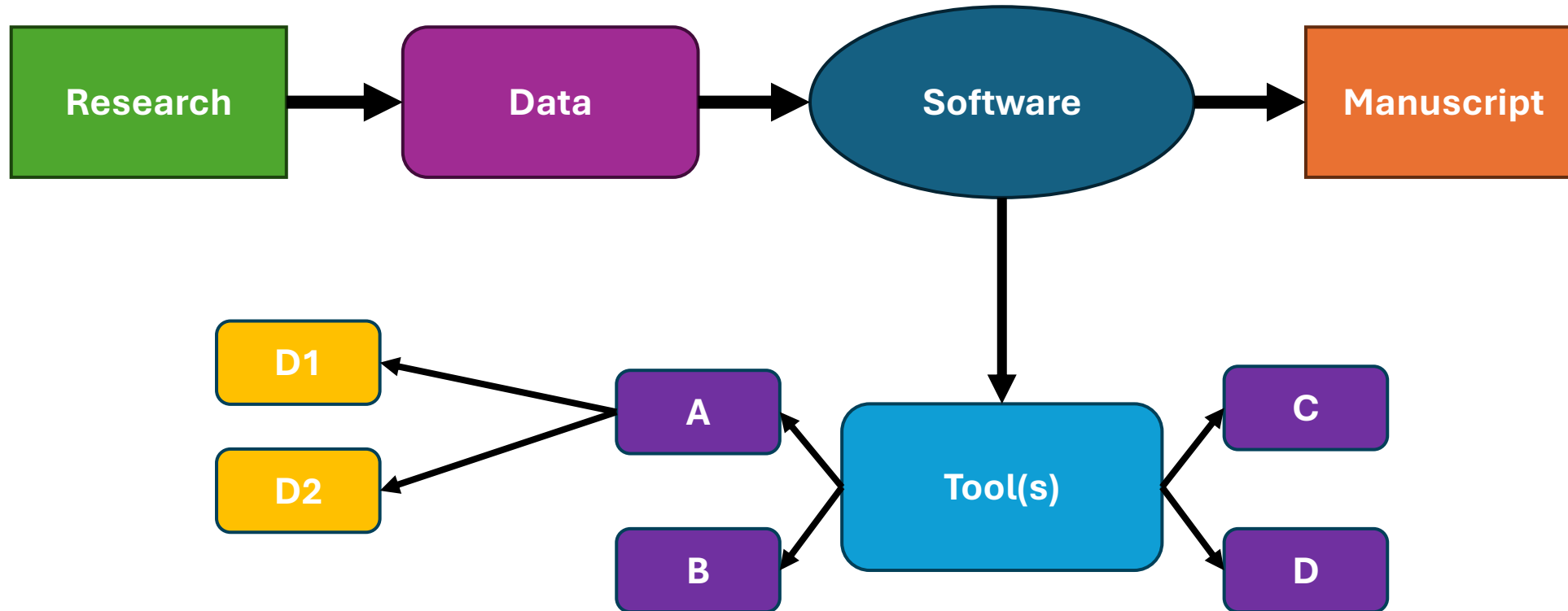
TOOLS &  
PIPELINES CAN  
EASILY BE  
SHARED &  
IMPLEMENTED.



EQUIVALENT TO  
DRAWERS IN  
FILING  
CABINET

# Dependency Management

YOU ARE HERE



INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS

# Conda Advantages

- Easy, fast and smooth installation of dependencies
- Isolated virtual environment like virtualenv in Python
- High versatility across various operation systems
- Significantly supports reproducibility outcomes
- Built-in support for a variety of programming languages

YOU ARE HERE

The logo for Conda, featuring a green circular icon with a white snake-like pattern on the left, followed by the word "CONDA" in a bold, green, sans-serif font.

INTRO

DEPENDENCY  
MANAGEMENT

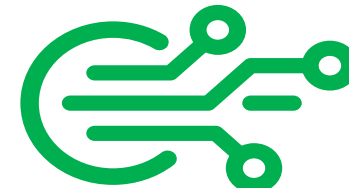
ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS

# Conda & Bioconda

- **Anaconda:** Original Scientific Python Distribution which includes Conda
- **Conda:** Useful Package Manager to manage dependencies
- **Miniconda:** installer package of conda
- **Bioconda:** channel of Bioinformatics packages for Conda



INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS

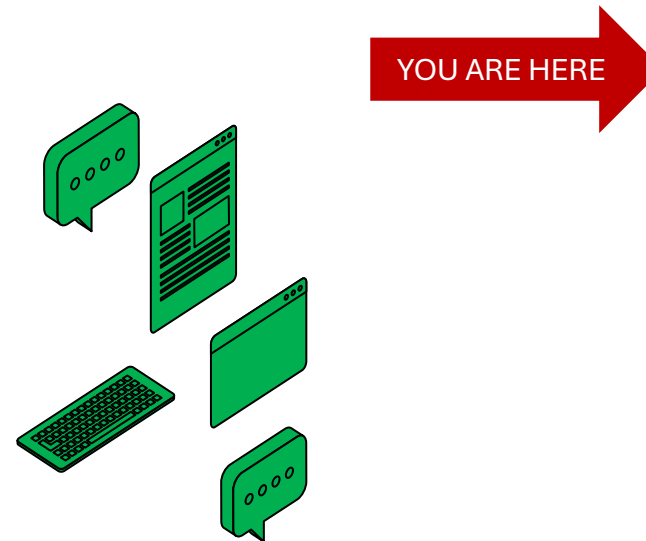
# Bioinformatics Software Packages

## •Bioinformatics Workflow:

- Over 80% of bioinformatics depends on proper software libraries management
- Traditional OS package managers (e.g., apt, dnf, Microsoft Store) lacks bioinformatics packages.

## •Mamba:

- A faster alternative to conda for solving dependencies & speeding up installations.



INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS



# Bioinformatics Software Packages

- **Channels:** Collections of packages used by conda and mamba.
- **Important Channels:**
  - **base Channel:** Originally created by Anaconda.org developers.
  - **conda-forge:** General-purpose and scientific software.
  - **bioconda:** Software specific to bioinformatics.
- Software is packaged for specific OS (e.g., Linux, Windows, macOS) and CPU architectures (e.g., x86, M1).
- Package managers install software compatible with the system's OS and CPU architecture.



INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS

# Mamba Installation

- **For Linux on x86\_64 / amd64, download using:**

```
wget -c https://github.com/conda-  
forge/miniforge/releases/latest/download/Mambaforge-Linux-x86_64.sh
```

```
bash Mambaforge-Linux-x86_64.sh
```

- **For MacOS on x86\_64 / amd64, download using:**

```
curl -O -J -L -C - https://github.com/conda-  
forge/miniforge/releases/latest/download/Mambaforge-MacOSX-x86_64.sh
```

```
bash Mambaforge-MacOSX-x86_64.sh
```



INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS

# Mamba Channel Setup

```
mamba config --add channels defaults
```

```
mamba config --add channels bioconda
```

```
mamba config --add channels conda-forge
```

```
mamba config --set channel priority strict
```

YOU ARE HERE



INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

DATA  
MANIPULATION

FILTERING FILE  
DATA

# Listing Packages

mamba list

# List all installed packages with conda list

mamba list ^z

# List all installed packages matching a regular expression with conda list {regex}

mamba search

# List all versions of all packages in all conda channels with conda search

mamba search scikit-learn

# List all versions of a package (all channels) with conda search {pkg}

mamba search 'scikit-learn>=1'

# List specific versions for a package (all channels) with conda search '{pkg}{version}'

mamba search conda-forge::scikit-learn

# List package versions in a specific conda channel with conda search {channel}::{pkg}

YOU ARE HERE

INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS

# Installing Packages

```
mamba install numpy pandas
```

```
# Install packages with conda install {pkg1}  
{pkg2} ...
```

```
mamba install scipy=1.10.1
```

```
# Install specific version of package with  
conda install {pkg}={version}
```

```
mamba update -all
```

```
# Update all packages with conda update --  
all
```

```
mamba uninstall pandas
```

```
# Uninstall packages with conda uninstall  
{pkg}
```

YOU ARE HERE

INTRO

DEPENDENCY  
MANAGEMENT

ADVANTAGES

CONDA &  
BIOINFORMATICS

CONDA/ MAMBA  
COMMANDS