

# The Carpentries Offline: Teaching Foundational Data Science and Coding Skills with little or no Internet Access



Abhishek Dasgupta / OxRSE Talk / 2024-02-23

# Team

**Jannetta Steyn** (Newcastle University)

**Abhishek Dasgupta** (University of Oxford)

**Colin Sauze** (National Oceanography Centre)

**Samantha Finnigan** (Durham University)

**Ethan White** (University of Florida)

**Virnaliz Cruz** (University of Florida)

**Frances Turner** (Newcastle University)

**Andrew Gill** (Stellenbosch University)



# THE CARPENTRIES

“ We teach foundational coding and data science skills to researchers worldwide. ”

- **Vision:** to be the leading inclusive community teaching data and coding skills.
- **Workshops:** Software Carpentry, Data Carpentry, and Library Carpentry
- **Roles:** Instructors, helpers, Trainers, Maintainers, Mentors, and Core Team
- **Technologies:** Web based course notes, etherpad for shared notes, Github, Jupyter Notebooks

<https://carpentries.org>

# What is CarpentriesOffline?

Enable working with Carpentries material  
entirely offline, with mirrors of PyPI and  
CRAN as well as all lesson content

# Why CarpentriesOffline?

Democratise access to Carpentries material in areas where Internet access is limited, such as due to electricity outages and expensive bandwidth







# How it all started

Software Sustainability Institute  
Collaborations Workshop Hackday 2021

Running workshops without Internet  
access

Use Raspberry Pi as an access point and  
web server

Hackday winner

SSI Fellowship 2022 (Janetta)



# offlinedatasci

Offlinedatasci mirrors installers and repositories to enable offline installation

```
pip install offlinedatasci  
offlinedatasci install all /install/path
```

<https://github.com/CarpentriesOffline/offlinedatasci>

# offlinedatasci

## What we mirror

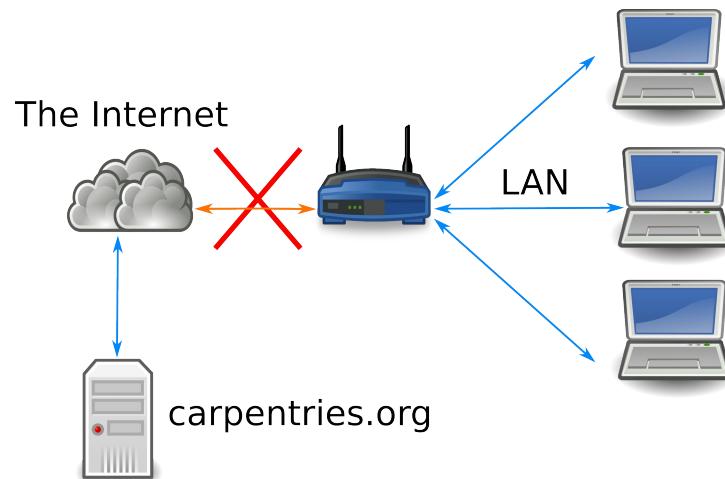
- Installers for Python and R
- Partial mirrors of PyPI, CRAN (packages can be customised)
- Carpentries online material
- Installers for data science IDEs (RStudio)

# Three threads to the project

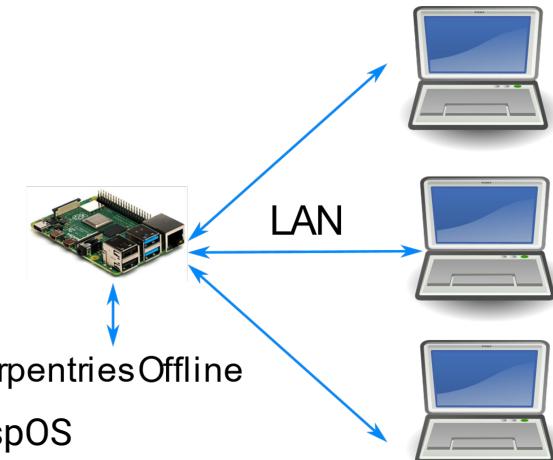
- Using a Raspberry Pi
- Using a bootable flash drive
- A mini HPC for HPC workshops



## The Problem



## The Solution



- RaspOS
- Gitea
- Etherpad
- Lesson Mirrors
- CRAN and PyPi mirrors



A composite screenshot of a computer desktop environment. At the top, a 'Wireless' status window shows network connections and battery levels: TALKTALK7D7AB3 (67%), carpentries-offline (84%), TALKTALK7D7AB3\_EXT (100%), and dvcpbcrccwd (47%). Below this is a dark-themed desktop with a taskbar at the bottom containing icons for File Explorer, Task View, Start, Task View, Network, and a search bar showing the IP address 192.168.1.1. A browser window is open to 'Welcome to Carpentries Offline', featuring the 'THE CARPENTRIES OFFLINE' logo and three main sections: 'Course Notes' (Data Carpentry, Software Carpentry), 'Services' (Gitea), and 'Downloads' (R, R Studio, Python).

# Carpentries Offline on the Pi

- Entire build process is scripted.
- Building images on the Raspberry Pi is a manual and slow process.
- Cloud based GitHub actions build in a Raspberry Pi emulator (Qemu).
  - Emulators are slow! Takes 2+ hours to build

# CarpentriesOffline in the Cloud

- Docker container using CarpentriesOffline build script
- Originally intended for testing.
- Much faster than using a Raspberry Pi or an emulator.
- Useful for when the Carpentries website/etherpad goes down during your workshop!
- Can be hosted in intranet

# Carpentries Offline on your laptop <alpha>

- RPis cost money
- I already have a laptop
- Bootable drive using Slax



# Carpentries Offline miniHPC <alpha>

- Hardware more visible
- Hit resource limitations more easily so more obvious
- No accounts to be setup on a real HPC
- No interfering with real HPC
  - users less afraid to try stuff
  - less likely to break anything important
  - no access to a real HPC
- Access problems
- Networks access

# miniHPC <alpha>

## Pixie the Prototype

- 3 x Raspberry Pi 4 B
- 1 x Raspberry Pi 4 B head/login node
- Raspberry Pi OS Lite (64 bit)(Debian Bullseye)
- Head node acts as WiFi access point

## RockPi

- 8 x Rock 4C+ (Dual ARM Cortex-72 @1.5GHz per node)
- 1 x Rock 4SE head/login node
- 8 x Power over Ethernet hats
- *Raxda* build of Debian Bullseye
- Head node acts as WiFi access

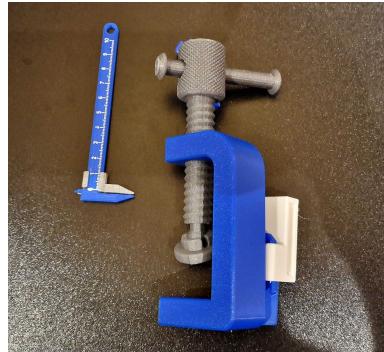


# HPC Software

- Slurm
- Lsmod
- Munge
- NFS
- PXE
- EasyBuild
- dnsmasq
  - DHCP
  - tftp
- mpich
- gcc
- python

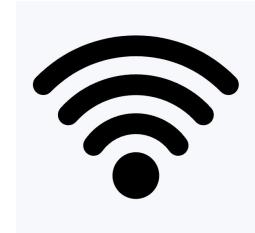
# 3D Printing Credit

- [https://www.printables.com/@TaylorSteinf\\_1252185](https://www.printables.com/@TaylorSteinf_1252185)
- <https://www.printables.com/model/717134-mini-caliper-10-cm>
- <https://www.printables.com/model/271563-printable-precision-measuring-tools>
- <https://www.thingiverse.com/thing:2424354>
- <https://www.printables.com/model/44122-customizable-belt-buckle>



# Find us online

- **Raspberry Pi Image download:**  
<https://github.com/carpentriesoffline/carpentriesoffline-installer/releases>
- **Find Us:**  
<https://carpentriesoffline.org>
- **Slack channel**  
<https://carpentries.slack.com/archives/C03KYQ3PX99>
- **Local URL:** <http://192.168.1.1>



carpentries-offline

