

# CFF & JSON for More Impactful RSE Training Materials

---

Toby Hodges  
The Carpentries

Jan Bernoth  
University of Potsdam

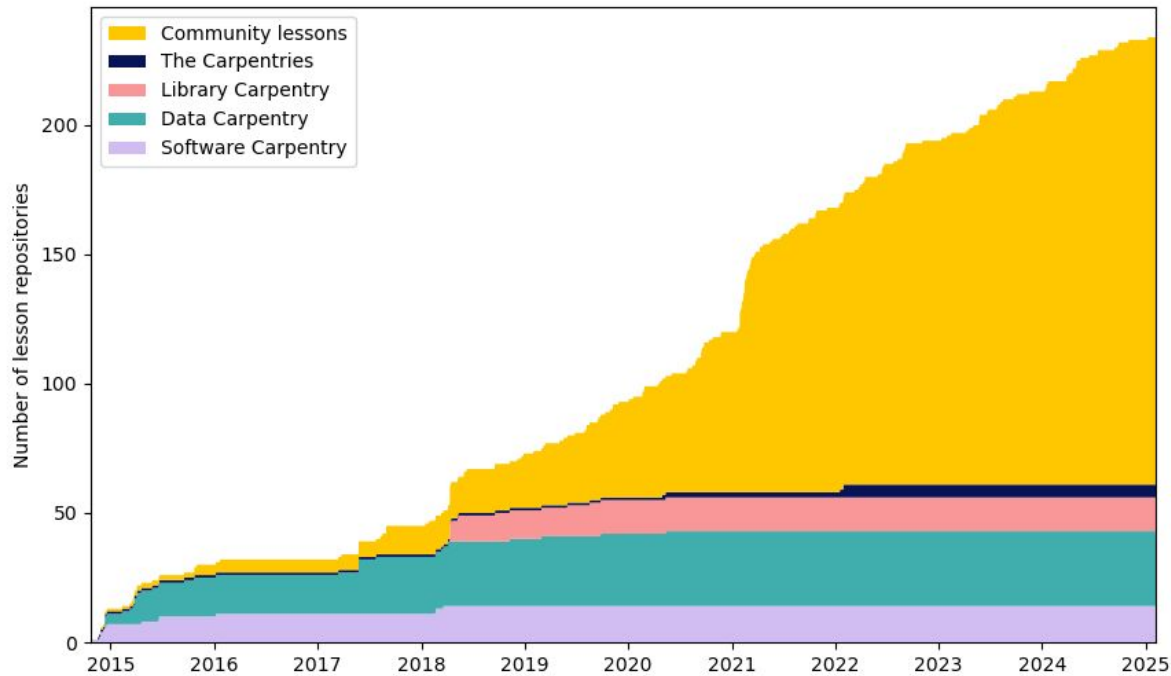


# The Carpentries

A global community teaching essential software and data skills.

- empower a diverse global community by equipping individuals with essential data and computational skills.
- promote efficient, open, and reproducible research practices.
- foster collaborative development of openly available lessons and research-informed teaching methods
- build an inclusive network of practitioners and advocates who emphasise the critical role of software and data in advancing research.

Three official lesson programs + >120 community-owned lessons.



The Carpentries community has >150 lessons in the Incubator + Lab

# RSEng-related lessons in the Incubator + Lab

- Intermediate RSEng with Python
  - Introduction to Containerised Computing with Docker|Singularity
  - Programming with Julia
  - Parallel Programming
  - GPU Programming
  - Workflow Management
  - Packaging and Publishing with R|Python
  - many Machine Learning + genAI lessons
  - ...
-

# Support for Lesson Development

## Infrastructure

- Open source, MIT licensed
- Write lessons with (R) Markdown
- Accessible: WCAG AA+
- Automated a11y testing
- Simple local setup

## Creating Histograms

Last updated on 2023-09-18 | [Edit this page](#)

[Expand All Solutions](#) +

### OVERVIEW

#### Questions

- How can we create grayscale and colour histograms to understand the distribution of colour values in an image?

#### Objectives

- Explain what a histogram is.
- Load an image in grayscale format.
- Create and display grayscale and colour histograms for entire images.

In this episode, we will learn how to use scikit-image

### First, import the packa

```
import imageio.v3 as iio
import ipyml
import matplotlib.pyplot as plt
import numpy as np
import skimage as ski

%matplotlib widget
```



#### HISTOGRAMS IN MATPLOTLIB

Matplotlib provides a dedicated function to compute and display histograms: `plt.hist()`. We will not use it in this lesson in order to understand how to calculate histograms in more detail. In practice, it is a good idea to use this function, because it visualises histograms more appropriately than `plt.plot()`. Here, you could use it by calling `plt.hist(image.flatten(), bins=256, range=(0, 1))` instead of `np.histogram()` and `plt.plot()` (`*.flatten()` is a NumPy function that converts our two-dimensional image into a one-dimensional array).



#### USING A MASK FOR A HISTOGRAM (15 MIN)

Looking at the histogram above, you will notice that there is a large number of very dark pixels, as indicated in the chart by the spike around the grayscale value 0.12. That is not so surprising, since the original image is mostly black background. What if we want to focus more closely on the leaf of the seedling? That is where a mask enters the picture!

First, hover over the plant seedling image with your mouse to determine the  $(x, y)$  coordinates of a bounding box around the leaf of the seedling. Then, using techniques from [the Drawing and Bitwise Operations episode](#), create a mask with a white rectangle covering that bounding box.

After you have created the mask, apply it to the input image before passing it to the `np.histogram` function.

Show me the solution



# Support for Lesson Development

The screenshot displays the 'The Carpentries Handbook' website. The header includes the site name and a search bar. A left sidebar contains a navigation menu with categories like 'Code of Conduct', 'GENERAL RESOURCES', and 'Lesson Development'. The main content area is titled 'Lesson Development / Lesson Pilot Workshops' and includes a 'Lesson Pilot Workshops' section with a 'Purpose' subsection. The 'Purpose' text describes the value of pilot workshops in lesson development. Below this is the 'Alpha and Beta Pilots' section, which outlines the stages of lesson development. At the bottom of the main content is the 'Information for Lesson Developers' section, which includes a subsection 'Finding Hosts for Beta Pilots'.

The Carpentries Handbook

THE CARPENTRIES

Search docs

Code of Conduct

GENERAL RESOURCES

Assessment

Communications

For Instructors

Fundraising

Governance

Instructor Development

Instructor Training

Lesson Development

Curriculum Advisory Committees

Lesson Infrastructure Subcommittee

Lesson Sprint Recommendations

Lesson Pilot Workshops

Purpose

Information for Lesson Developers

Information for Hosts

Release Process and Schedule

Lesson Release Checklist

The Carpentries Incubator Lesson Spotlight

Lesson Maintenance

Policies

Regional Communities

Teaching and Hosting

Workshop Administration

Lesson Development / Lesson Pilot Workshops

Edit on GitHub

## Lesson Pilot Workshops

### Purpose

Teaching a lesson for the first time is very rewarding, but the experience of the Instructors and learners also identifies opportunities to address and further clarify parts of the content. This makes these early runs through a lesson, which we refer to as *lesson pilots*, crucial milestones in the development of a high-quality lesson. As well as teaching new and exciting skills to learners, the additional purpose of pilot workshops is to collect information and feedback that can be used to polish content and make the lesson more reusable by other Instructors (e.g. by recording accurate timings for episodes and exercises, expanding Instructor Notes, etc.).

### Alpha and Beta Pilots

The lesson development process includes pilot workshops at two different stages, which we refer to as *alpha* and *beta* pilots. Alpha pilots are the first workshops where the lesson is taught, almost always by some or all of the original developers of the lesson.

After the feedback from these alpha pilots has been used to improve the lesson, it can enter the beta stage, where other Instructors - who did not have a major part in the previous development of the lesson - teach it and provide feedback.

For more information about these pilots, and the requirements for piloting official Carpentries lessons, see [the Lesson Life Cycle chapter of The Carpentries Curriculum Development Handbook](#).

### Information for Lesson Developers

#### Finding Hosts for Beta Pilots

If you are developing a new official Carpentries lesson - a lesson developed based on prior agreement with The Carpentries, and which is intended to become another lesson/curriculum offered in centrally-organised workshops - the Curriculum Team will help you find hosts and Instructors for pilot workshops.

## Communications

- Lesson & issue listings
- Blog posts
- Slack channel
- Mailing list

## Resources

- Lesson Developer Handbook
- Documentation
- Templates

## Training...



THE  
CARPENTRIES



**COLLABORATIVE  
LESSON DEVELOPMENT  
TRAINING**

# Lesson Metadata

## TrainingMaterial

- id, name description
- date created/modified/published
- url
- keywords, language



## Creating Histograms

Last updated on 2023-09-18 | [Edit this page](#)

[Expand All Solutions](#) +

### OVERVIEW

#### Questions

- How can we create grayscale and colour histograms to understand the distribution of colour values in an image?

#### Objectives

- Explain what a histogram is.
- Load an image in grayscale format.
- Create and display grayscale and colour histograms for entire images.
- Create and display grayscale and colour histograms for certain areas of images, via masks.

In this episode, we will learn how to use scikit-image functions to create and display histograms for images.

### First, import the packages needed for this episode

[PYTHON](#) < >

```
import imageio.v3 as iio
import ipynb
import matplotlib.pyplot as plt
import numpy as np
import skimage as ski

%matplotlib widget
```

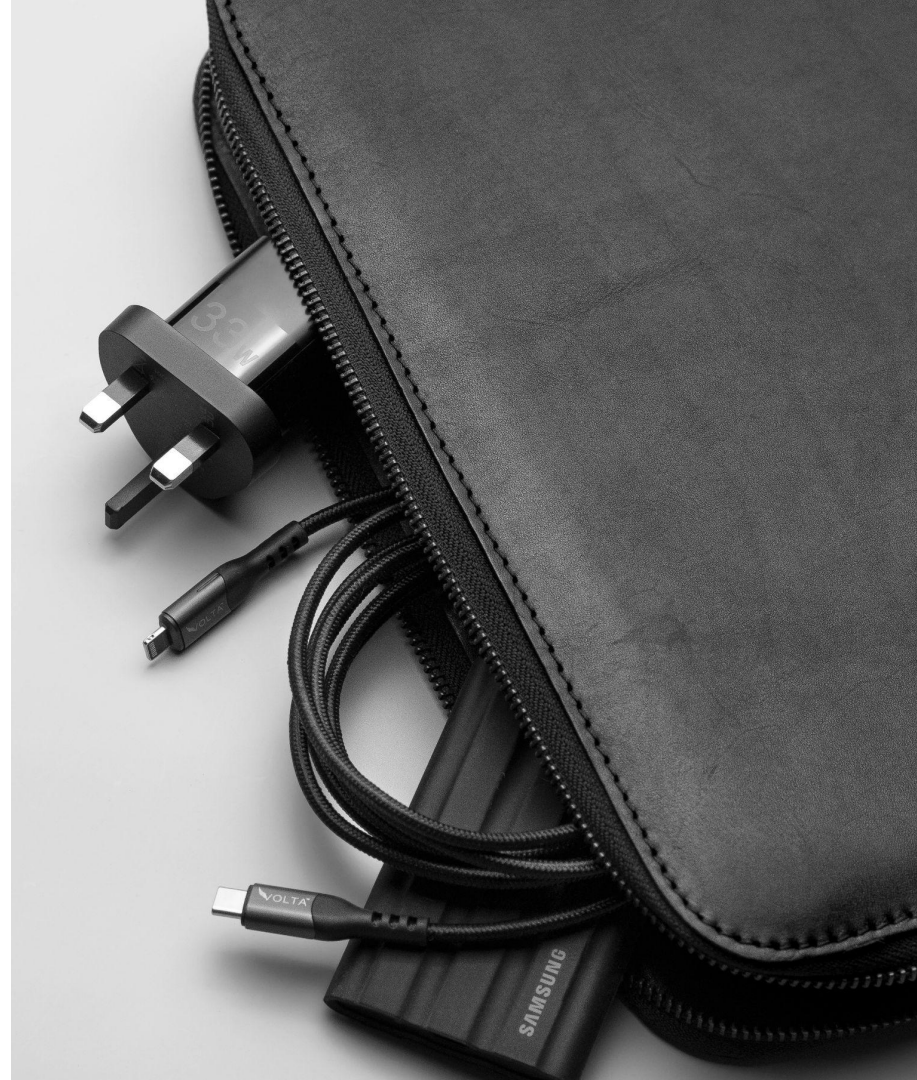


# Why Care About Metadata?

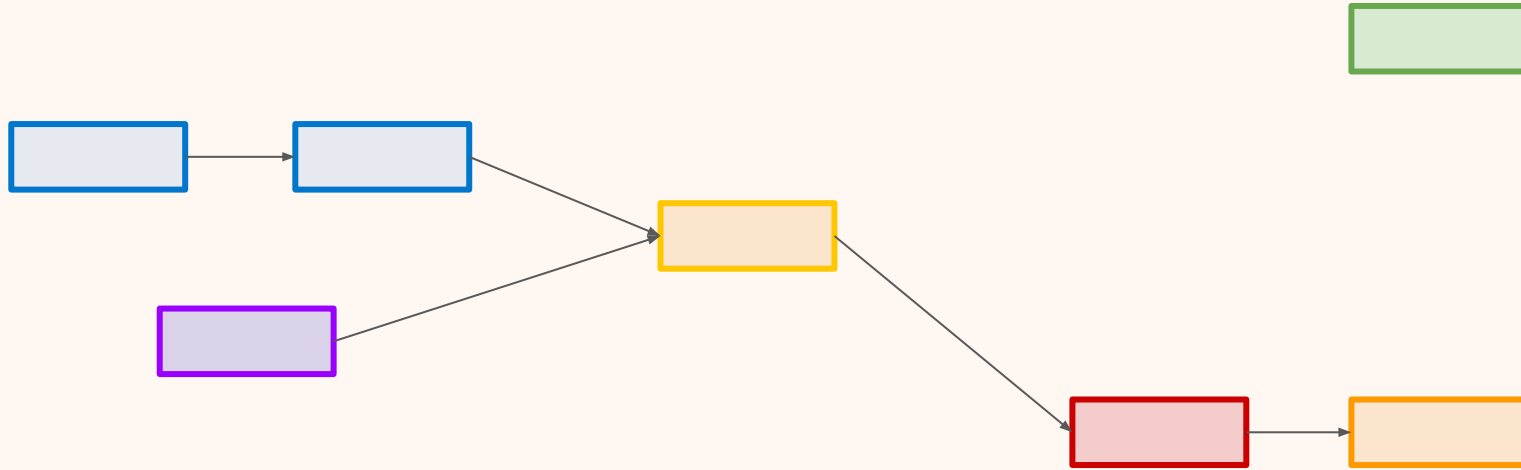
---

Lessons with **good** metadata are more findable and more interoperable.

Image by Sayan Majhi on Unsplash

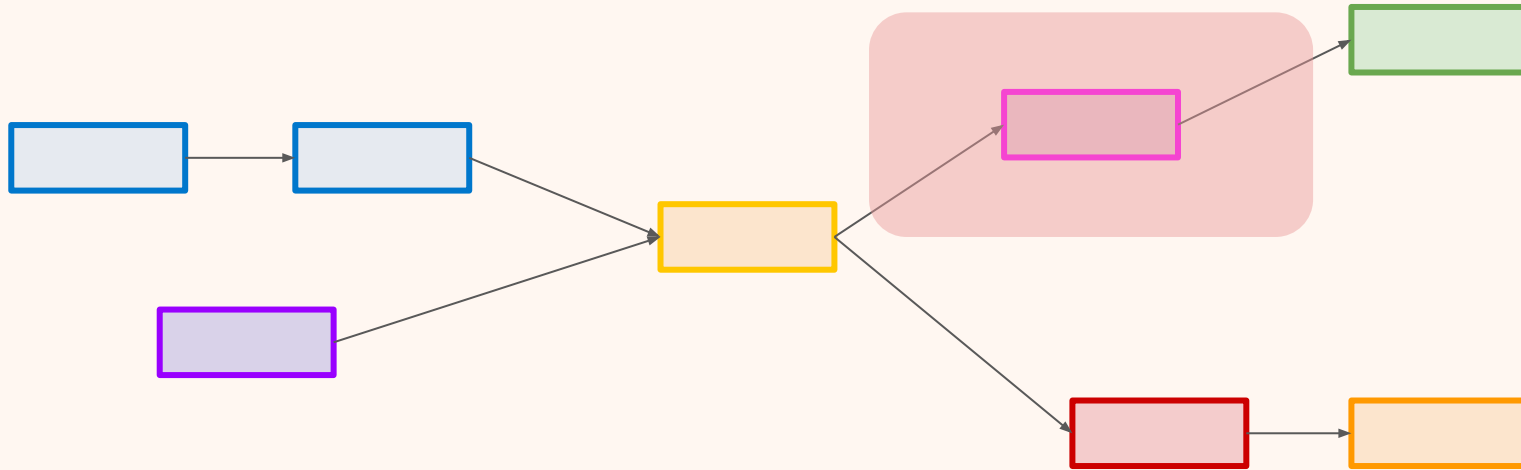


For example...



We could automatically present pathways through available resources...

For example...



... and identify gaps in these pathways

We should standardise how  
we describe the skills we  
want to teach and learn.

---

# CFFs for Lessons

---

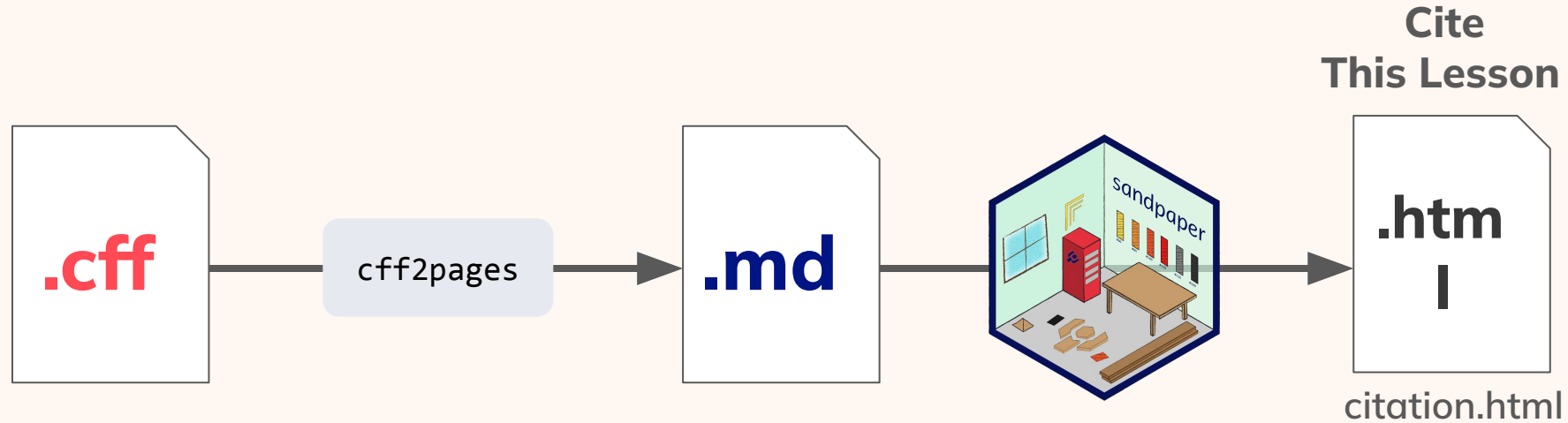
# CFF can provide some metadata currently missing...

|               |                    |
|---------------|--------------------|
| abstract      | license            |
| authors       | message            |
| cff-version   | preferred-citation |
| contact       | references         |
| date-released | repository-code    |
| doi           | title              |
| identifiers   | type               |
| keywords      | version            |

|                    |               |
|--------------------|---------------|
| description        | contributor   |
| keywords           | identifier    |
| name               | inLanguage    |
| about              | datePublished |
| abstract           | license       |
| audience           | mentions      |
| author             | teaches       |
| competencyRequired | version       |

...and enable *Cite This Lesson* pages!

# Integrating CFF into The Carpentries Workbench



---

Please visit Poster 101! ***cff2pages: Expanding Workflows with Markdown Export*** — Jan Bernoth

Thanks for listening!

