

# **“YOU HAVE 0 CREDIT – PLEASE INSERT ~~COIN~~ FILE”. THE CITATION FILE FORMAT**

Stephan Druskat

Third Conference of Research Software Engineers, University of Birmingham,  
Birmingham, UK, 4 September 2018.

Slides: <https://sdruskat.net/rse18-citation-file-format/>

# WHO I AM

- RSE (linguistics, digital humanities) at Humboldt-Universität zu Berlin
- **FORCE11 Software Citation Implementation Working Group**

ORCID: 0000-0003-4925-7248

Email: [stephan.druskat@hu-berlin.de](mailto:stephan.druskat@hu-berlin.de)

Twitter: [@stdruskat](https://twitter.com/stdruskat)

GitHub: [@sdruskat](https://github.com/sdruskat)

**- WHAT IS THE CITATION FILE FORMAT  
(CFF)?**

**- WHY SHOULD YOU USE IT?**

# **WHAT IS CFF?**

## **A FORMAT FOR CITATION FILES**

**(CITATION files contain software citation metadata)**

# CITATION FILES - WHAT FOR?

- Papers usually contain their own citation metadata
  - *Software usually doesn't*

# SOLUTION

**SOFTWARE MUST PROVIDE ITS OWN CITATION METADATA**

- How? *In a CITATION file!*
- Where? *In the source code repository root!*

Wilson, Robin: "Encouraging citation of software – introducing CITATION files". SSI blog. <http://bit.ly/SSI-blog-citation-files>.

# Encouraging citation of software – introducing CITATION files

By [Robin Wilson](#), [Fellow](#) and postgraduate at the University of Southampton.

**Put a plaintext file named CITATION in the root directory of your code, and put information in it about how to cite your software. Go on, do it now: it'll only take two minutes!**

Software is very important in science – but good software takes time and effort that could be used to do other work instead. I believe that it is important to do this work, but to make it worthwhile, people need to credit for their work, and in academia that means citations. However, it is often very hidden away somewhere in the manual, the author asking them how they want it cited, or to cite the software they use, and thus not doing this, so that software – which underlies a huge amount of research – deserves.

As with many things relating to software sustainability in science, the R project does this very well: if you want to find out how to cite the R software itself you simply run the command:

```
citation()
```

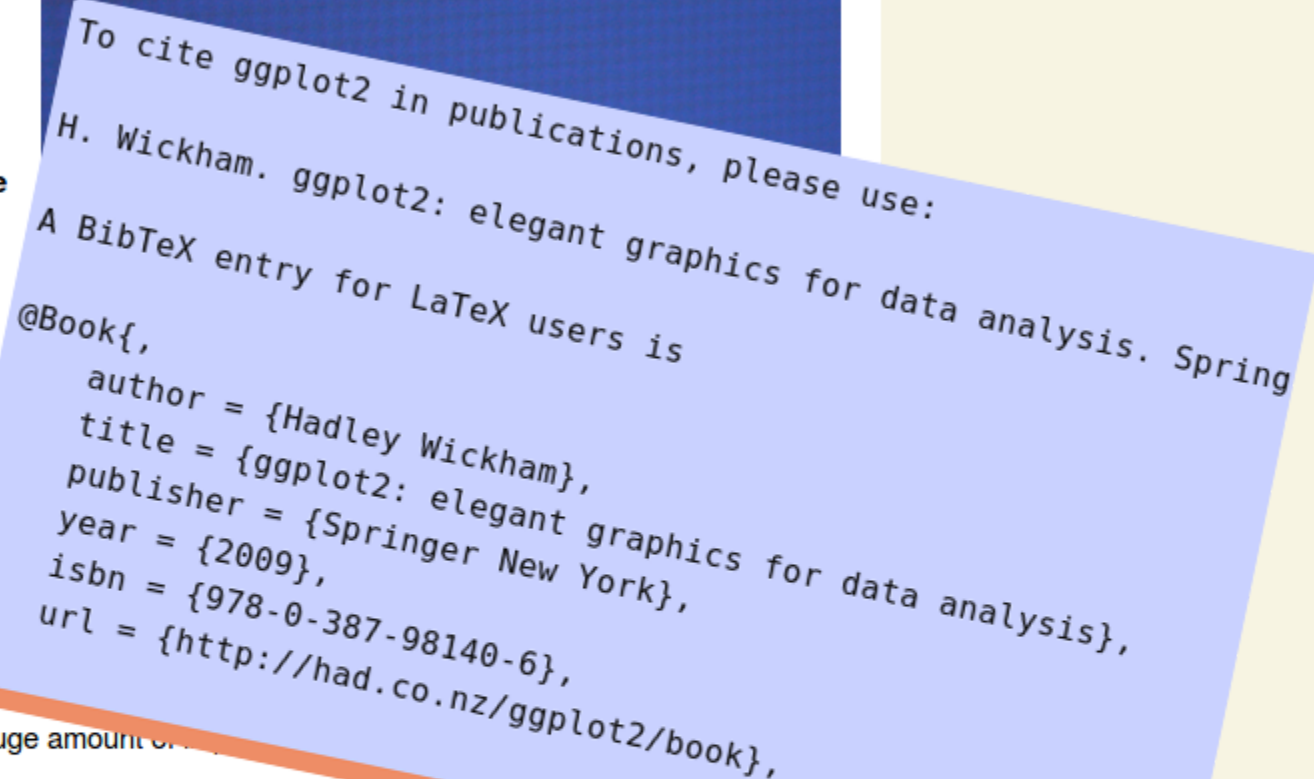
If you want to find out how to cite a package you simply run:

```
citation(PROJECTNAME)
```

For example:

```
> citation('ggplot2')
```

To cite ggplot2 in publications, please use:



```
To cite ggplot2 in publications, please use:  
H. Wickham. ggplot2: elegant graphics for data analysis. Springer  
A BibTeX entry for LaTeX users is  
@Book{,  
  author = {Hadley Wickham},  
  title = {ggplot2: elegant graphics for data analysis},  
  publisher = {Springer New York},  
  year = {2009},  
  isbn = {978-0-387-98140-6},  
  url = {http://had.co.nz/ggplot2/book},
```



To cite ggplot2 in publications, please use:

H. Wickham. ggplot2: elegant graphics for data analysis. Springer New York, 2009.

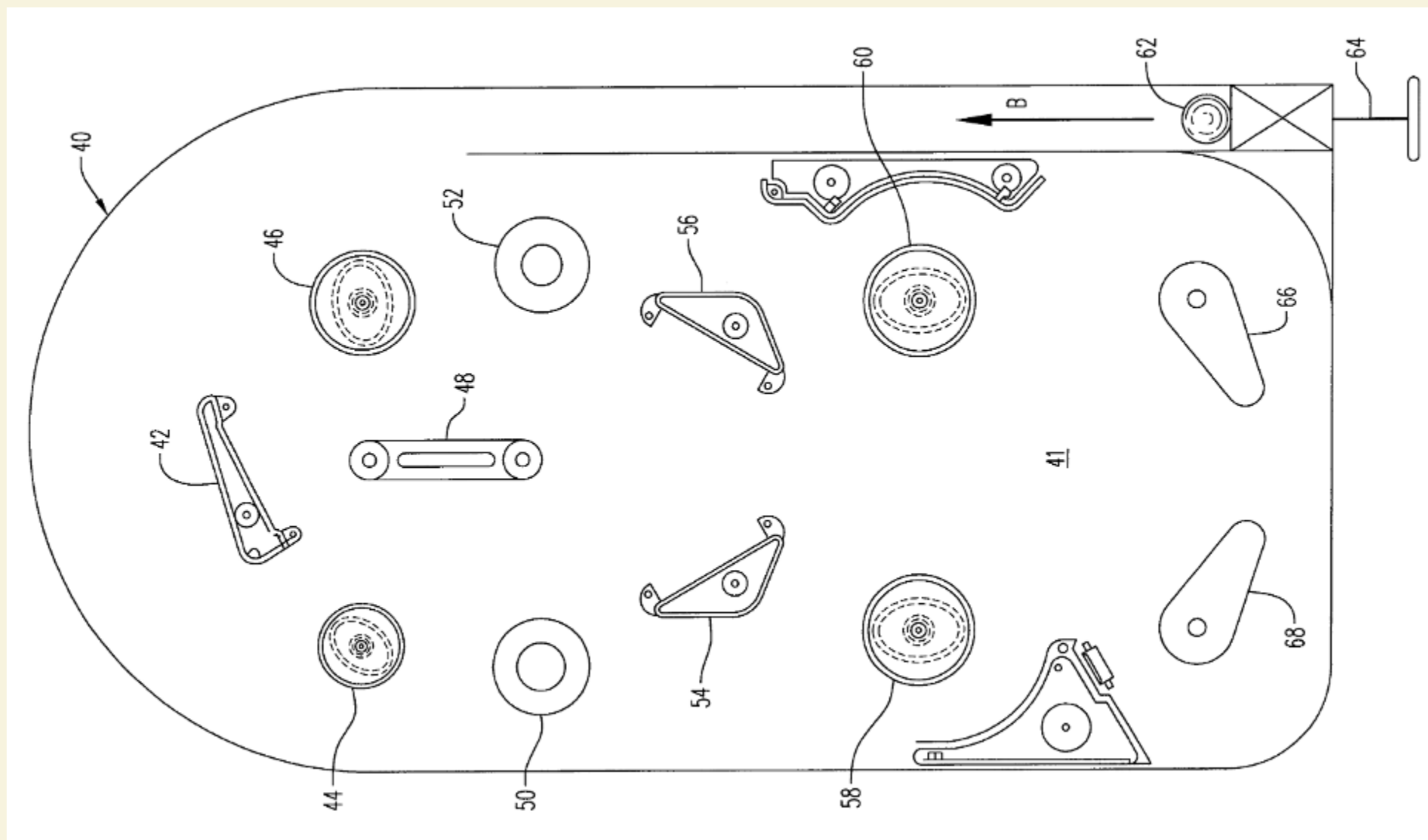
A BibTeX entry for LaTeX users is

```
@Book{,  
  author = {Hadley Wickham},  
  ...  
}
```

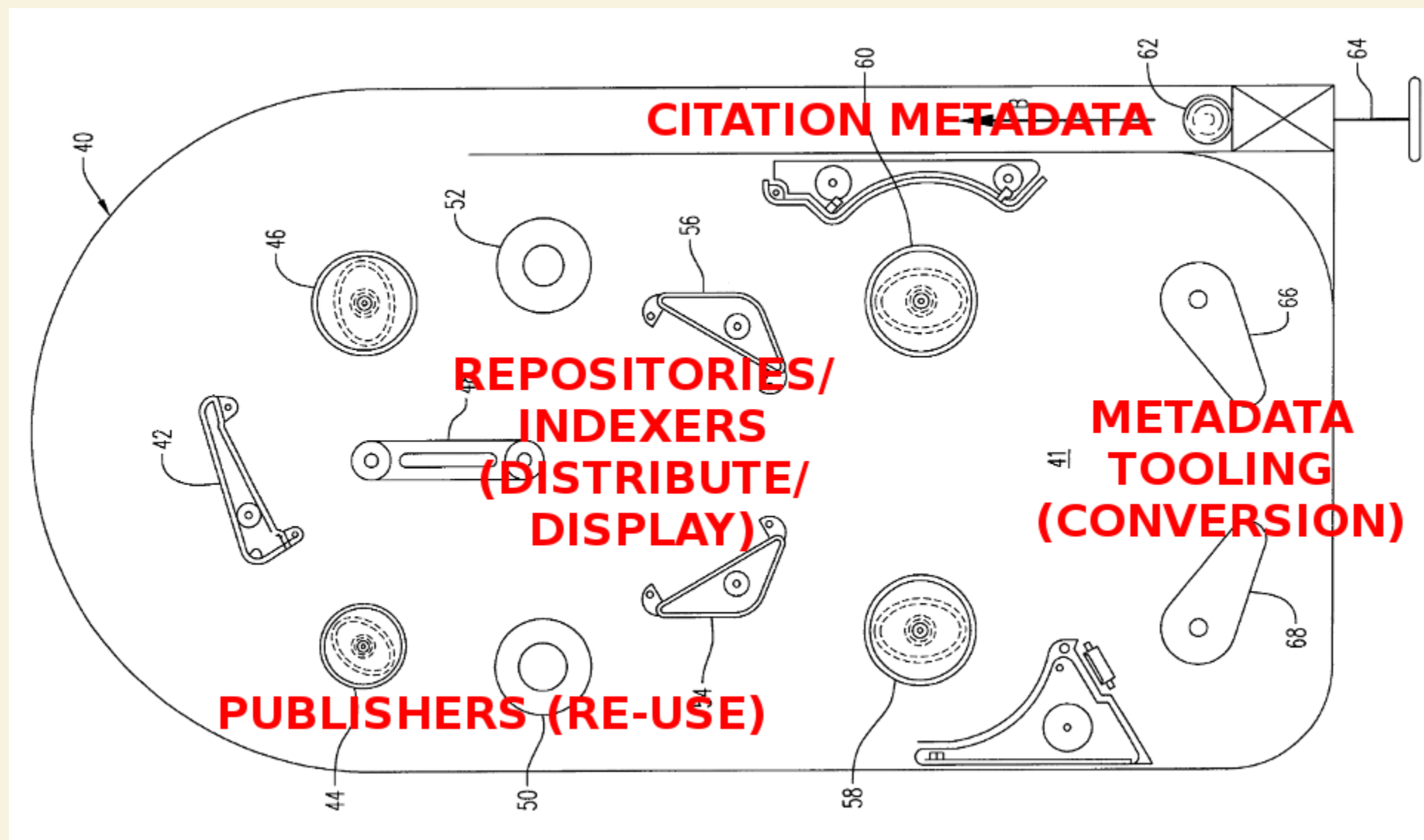
**AND THAT'S ALL YOU NEED!**  
**(IF YOU'RE A HUMAN)**  
**(AND WANT TO "JUST CITE A SOFTWARE")**

**SOFTWARE CITATION ISN'T JUST HUMANS  
CITING SOFTWARE IN PAPERS**

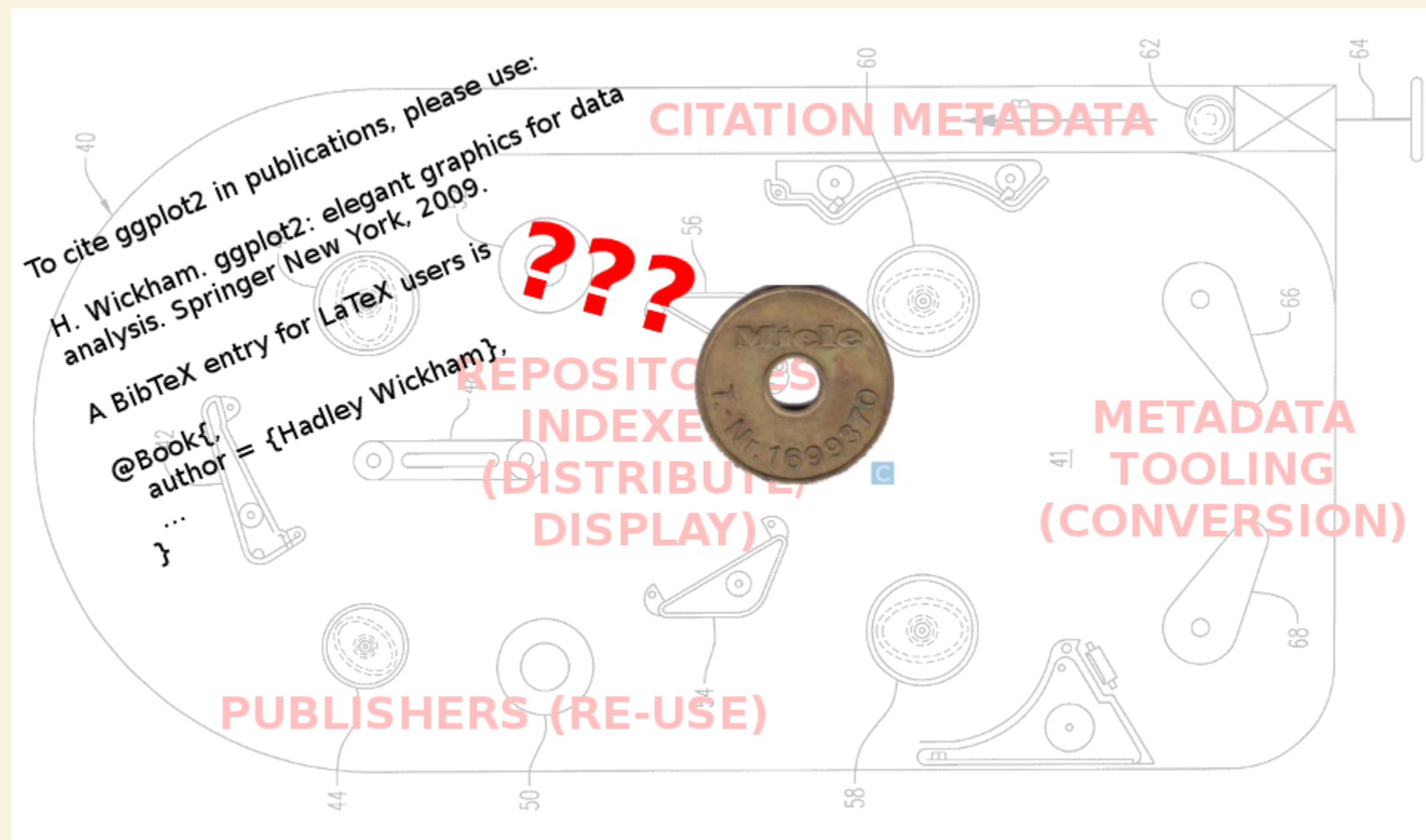
# SOFTWARE CITATION WORKFLOW



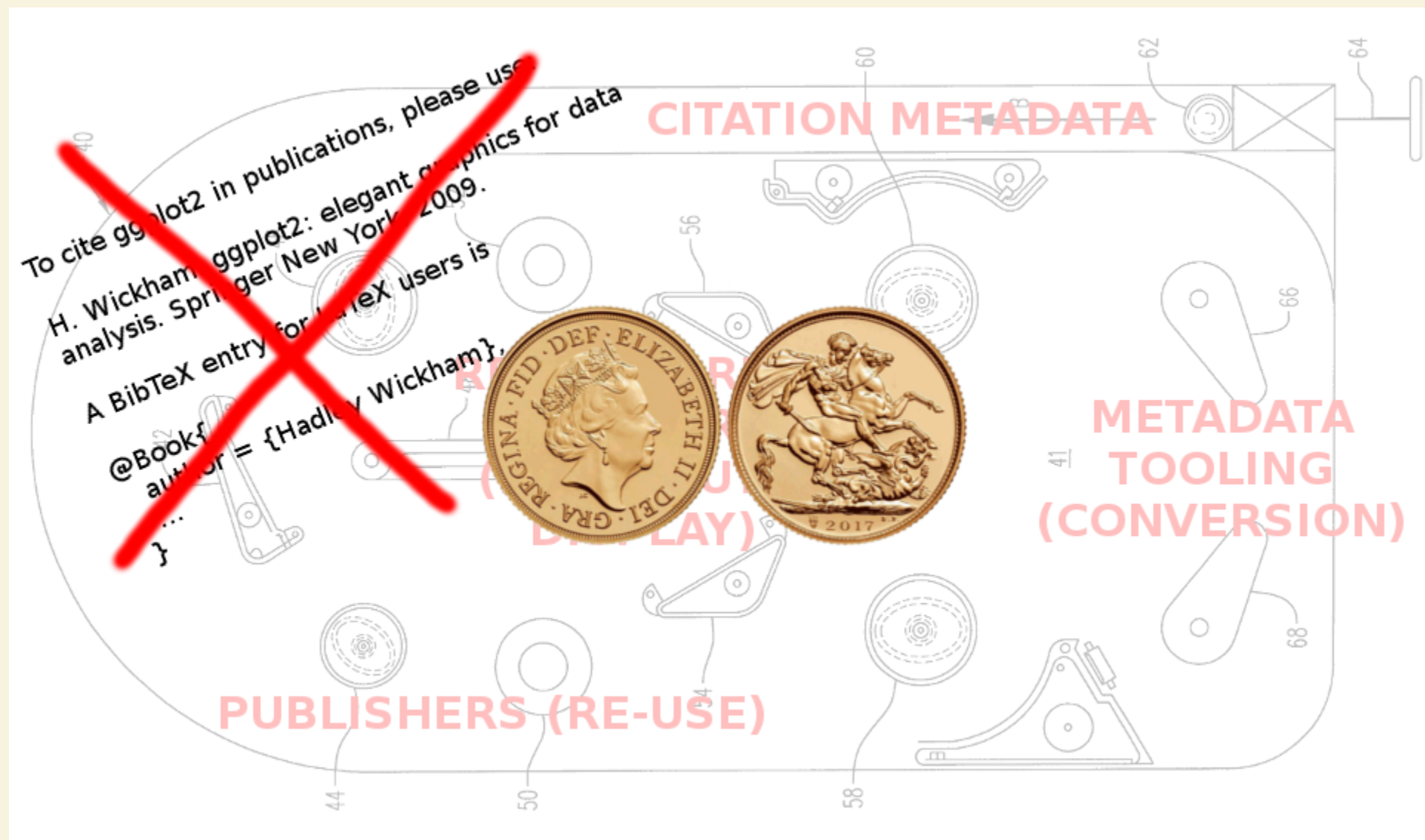
# SOFTWARE CITATION WORKFLOW



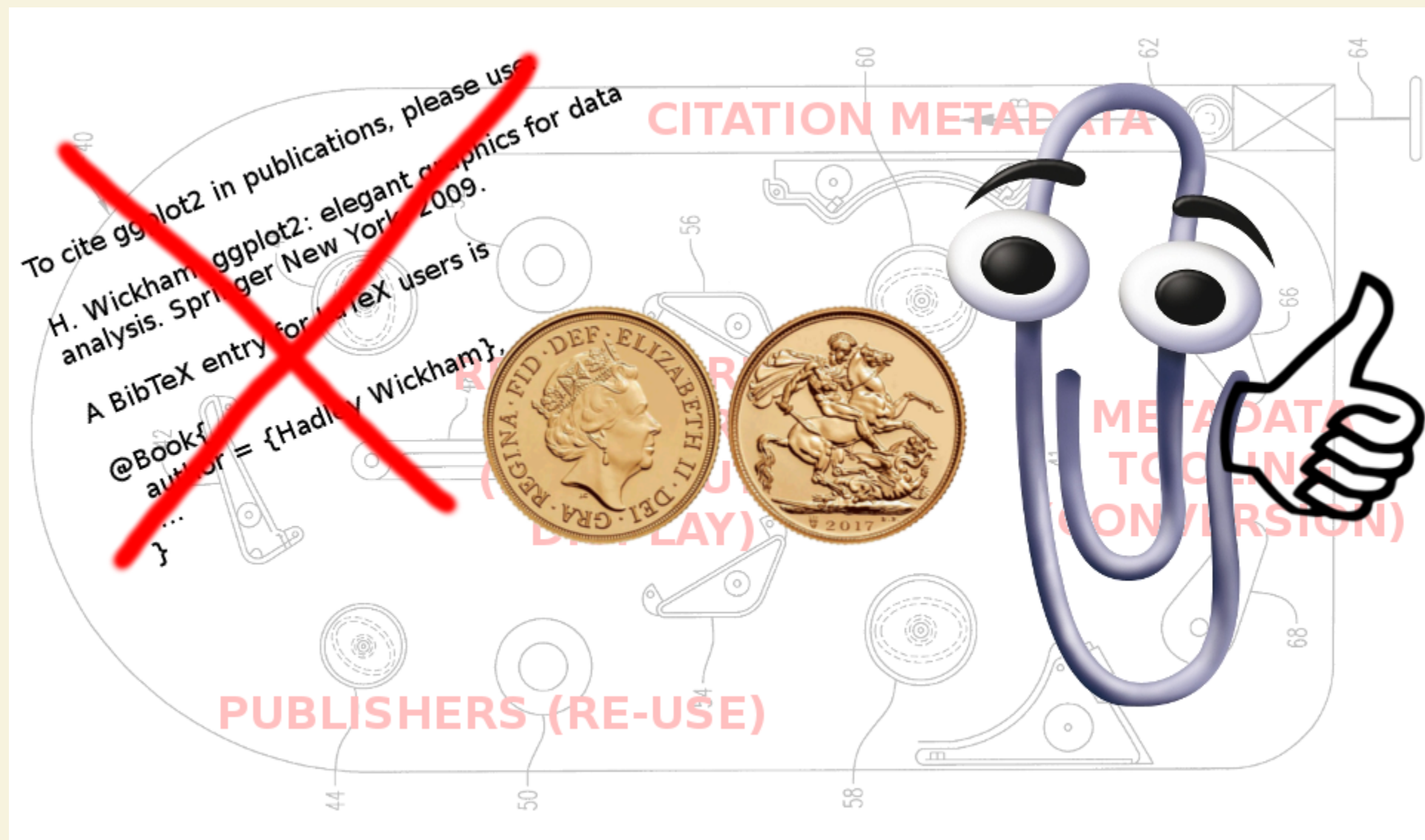
# SOFTWARE CITATION WORKFLOW



# SOFTWARE CITATION WORKFLOW



# SOFTWARE CITATION WORKFLOW





**WHAT IS CFF *EXACTLY*?**

***A machine-readable* format for CITATION files**

# CITATION.cff

```
cff-version: 1.0.3
message: "If you use this software, cite it as below."
authors:
  - family-names: Druskat
    given-names: Stephan
    orcid: "https://orcid.org/0000-0003-4925-7248"
title: "My Research Tool"
version: 1.0.4
doi: "10.5281/zenodo.1234"
date-released: 2017-12-18
```

# CITATION.cff

```
cff-version: 1.0.3
message: "If you use this software, cite it as below."
authors:
  - family-names: Druskat
    given-names: Stephan
    orcid: "https://orcid.org/0000-0003-4925-7248"
    affiliation: "Humboldt Berlin, Dept. of ..."
    email: mail@sdruskat.net
    website: https://hu.berlin/sdruskat
title: My Research Tool
...
```

# CITATION.cff

```
cff-version: 1.0.3
message: "If you use this software, cite it as below."
...
repository-code: https://github.com/sdruskat/mrt
repository-artifact: https://hu.berlin/nexus/mrt
license: Apache-2.0
url: https://sdruskat.github.io/my-research-tool
keywords:
  - "McAuthor's algorithm"
  - deeply confusional neural network
```

```
message: "Please cite MRT AND the outline paper." ...
references:
- type: article
  scope: Cite this for general concepts of MRT.
  authors:
    - family-names: Doe
      given-names: Jane
  title: "A 100% accuracy parser for all languages"
  year: 2099
  journal: Journal of Hard Science Fiction
  volume: 42
  issue: "13"
  doi: "10.9999/hardscifi-lang.42132"
```

# WHY I WANT YOU TO TRY CFF

- CFF enables correct citation of software
- CFF enforces **Software Citation Principles**
  - CFF is human-readable and -writable
- CFF is self-documenting (CITATION.cff, message)
  - CFF is compatible with downstream formats

# HOW TO USE CFF

1. Copy and paste the example from the **CFF homepage**
2. Replace details with your own

# TOOLING

- Getting started: **CFF initializer** (Jurriaan Spaaks)
- **Tools for conversion, creation, manipulation, etc.** in different programming languages available: <https://github.com/citation-file-format/>
- **Software Citation Central Station (SoCCS)**, development in progress, check back after the CFF Hack Day on Wednesday 5 Sep



# WHAT HAVE WE LEARNED?

1. **Machine-readable CITATION files are needed to harness the software citation workflow from the start**
2. **The Citation File Format is suitable, and provides further features**
3. **You should try it NOW and add a file to your software**
4. **It's got nothing to do with pinball really**

# THANKS!

**Contributors:** James W. Baker, Radovan Bast, Neil Chue Hong, Michael R. Crusoe, Morane Gruenpeter, Robert Haines, Tom Klaver, Alexander Konovalov, David McKain, Joseph T. Parker, Andrew Rowley, Ranieri Silva, Jurriaan H. Spaaks, Stefan Verhoeven

**Early CFF adopters:** The Netherlands eScience Center

**RSE18 Mentoring Programme:** Neil Chue Hong

**The Software Sustainability Institute** for funding assistance (EPSRC, BBSRC, ESRC Grant EP/N006410/1)



# RESOURCES

- Citation File Format website: <https://citation-file-format.github.io/>
- Citation File Format GitHub org: <https://github.com/citation-file-format/>