

DANIEL FALSTER, NICK GOLDING, SARAS WINDECKER, & JIAN YEN

TAKING YOUR R SKILLS TO THE NEXT LEVEL

TODAY WE ARE MEETING ON LUTRUWITA (TASMANIA)
ABORIGINAL LAND, SEA AND WATERWAYS. WE
ACKNOWLEDGE, WITH DEEP RESPECT THE
TRADITIONAL OWNERS OF THIS LAND, THE PALAWA
PEOPLE, WHICH WE MEET TODAY.

WE ARE ADOPTING THE ESA CONFERENCE CODE OF CONDUCT. THE COC IS DESIGNED TO CREATE INCLUSIVENESS, MAKE PEOPLE FEEL COMFORTABLE TO ASK QUESTIONS AND PROVIDE A SAFE ENVIRONMENT FOR PEOPLE OF ALL GENDERS, BELIEFS OR RELIGIONS.



©2014 DWA LLC. All Rights Reserved.

TINY.CC/ESA-QERC

NATURE | NEWS FEATURE

1,500 scientists lift the lid on reproducibility

Survey sheds light on the ‘crisis’ rocking research.

Monya Baker

25 May 2016 | Corrected: 28 July 2016

NATURE | NEWS FEATURE

1,500 scientists lift the lid on reproducibility

Survey sheds light on the ‘crisis’ rocking research.

Monya Baker

25 May 2016

THE CONVERSATION

Academic rigour, journalistic flair

Arts + Culture Business + Economy Cities Education Environment + Energy Health + Medicine Politics +

The science ‘reproducibility crisis’ – and what can be done about it

March 15, 2017 8.49pm AEDT

IRREPRODUCIBLE RESULTS... WHY?

Fidler F, Chee YE, Wintle BC, Burgman MA, McCarthy MA, Gordon A. Metaresearch for Evaluating Reproducibility in Ecology and Evolution. *Bioscience*. 2017;67(3):282-289. doi:10.1093/biosci/biw159

IRREPRODUCIBLE RESULTS... WHY?

1. PUBLICATION BIAS

IRREPRODUCIBLE RESULTS... WHY?

1. PUBLICATION BIAS
2. QUESTIONABLE RESEARCH PRACTICES
IN 'PUBLISH OR PERISH' CULTURE

IRREPRODUCIBLE RESULTS... WHY?

1. PUBLICATION BIAS
2. QUESTIONABLE RESEARCH PRACTICES
IN 'PUBLISH OR PERISH' CULTURE
3. INADEQUATE DATA REPORTING

IRREPRODUCIBLE RESULTS... WHY?

1. PUBLICATION BIAS
2. QUESTIONABLE RESEARCH PRACTICES
IN 'PUBLISH OR PERISH' CULTURE
3. INADEQUATE DATA REPORTING
4. INSUFFICIENT INCENTIVES FOR
SHARING CODE & DATA

NATURE | NEWS FEATURE

1,500 scientists lift the lid on reproducibility

Survey sheds light on the ‘crisis’ rocking research.

Monya Baker

THE CONVERSATION

25 May 2016

Corrected: 28 July 2016

Arts + Culture Business + Economy Cities Education Environment + Energy Health + Medicine Politics +

nature
ecology & evolution

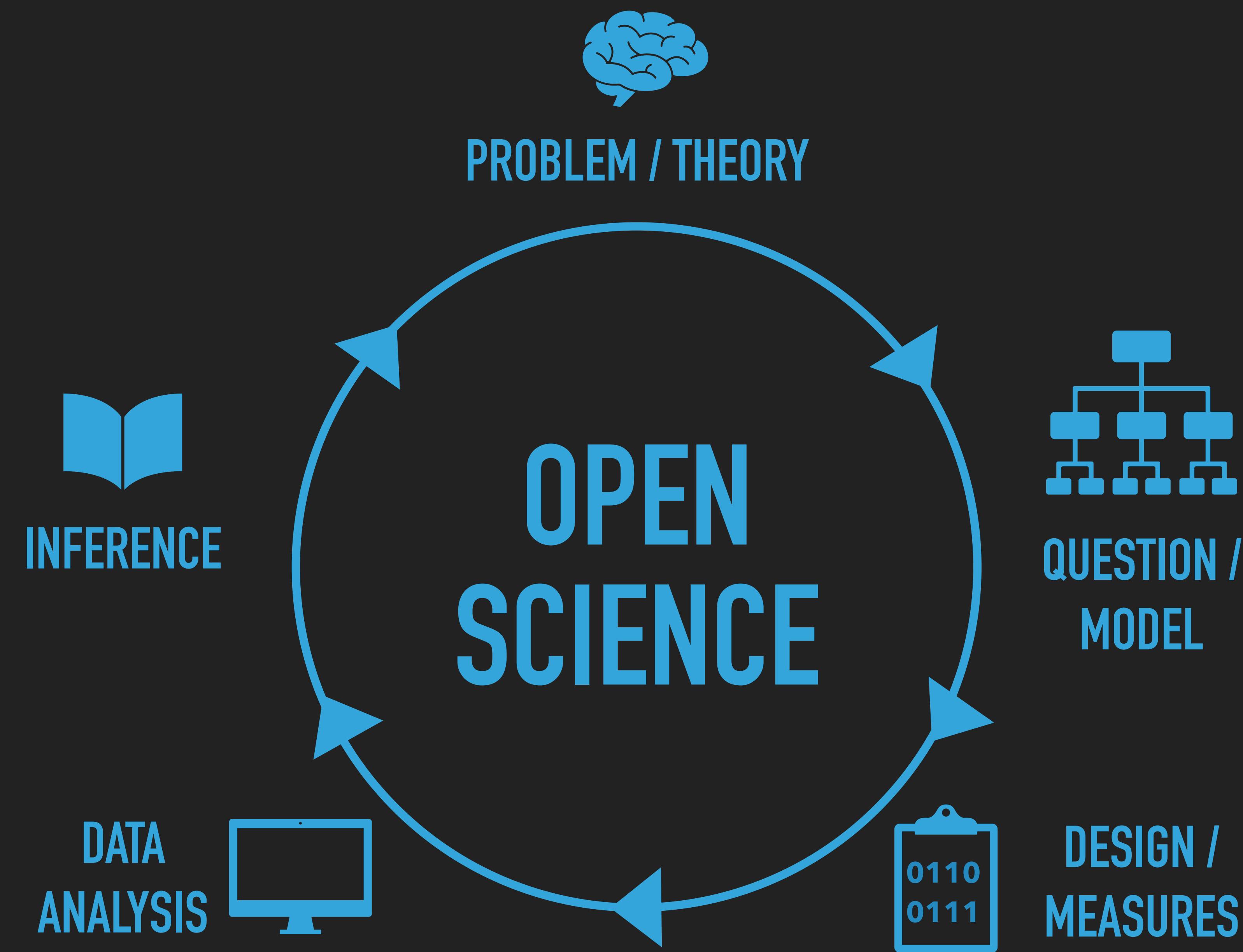
The sc
what c

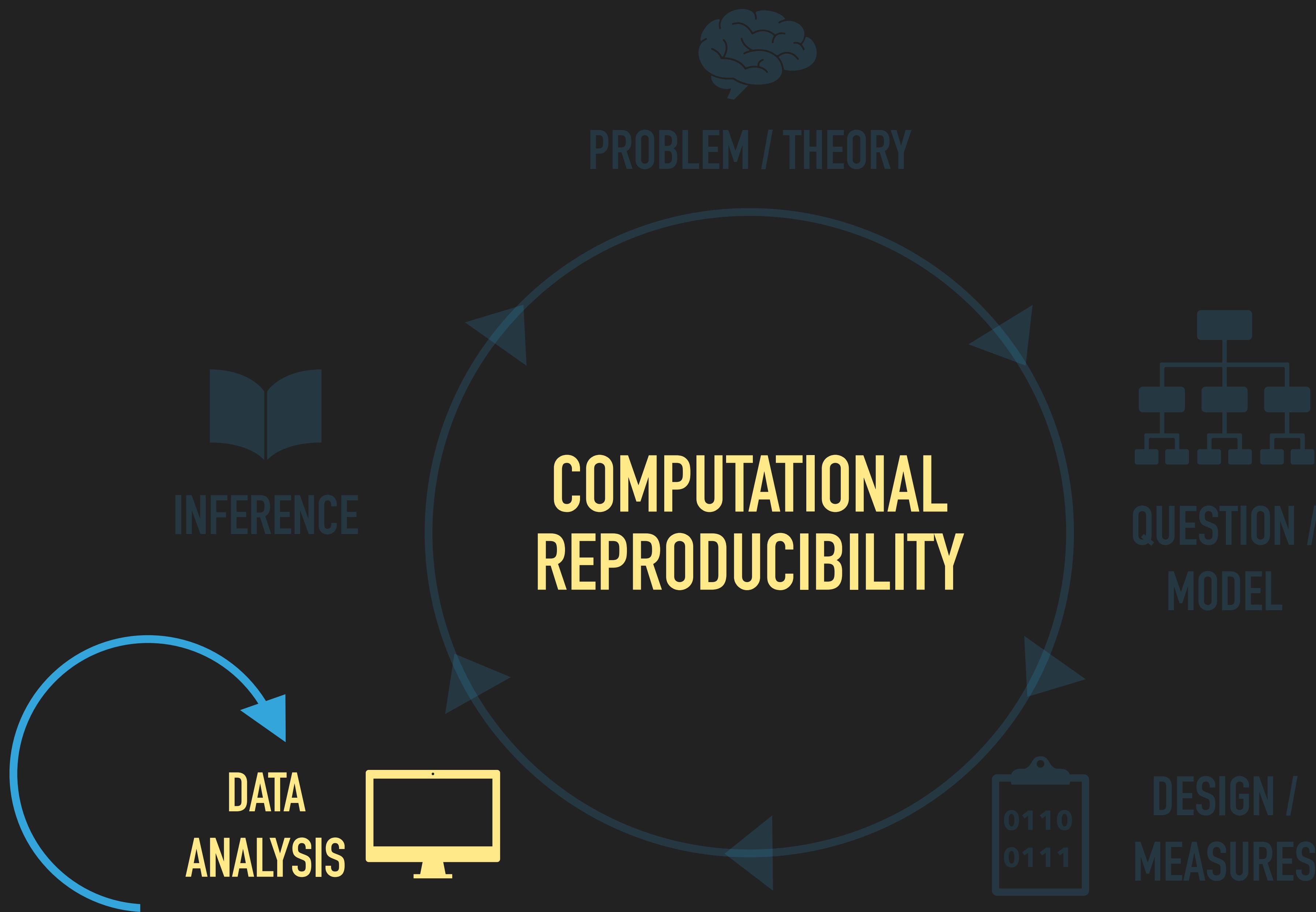
Perspective | Published: 23 May 2017

March 15, 2017 8.49p

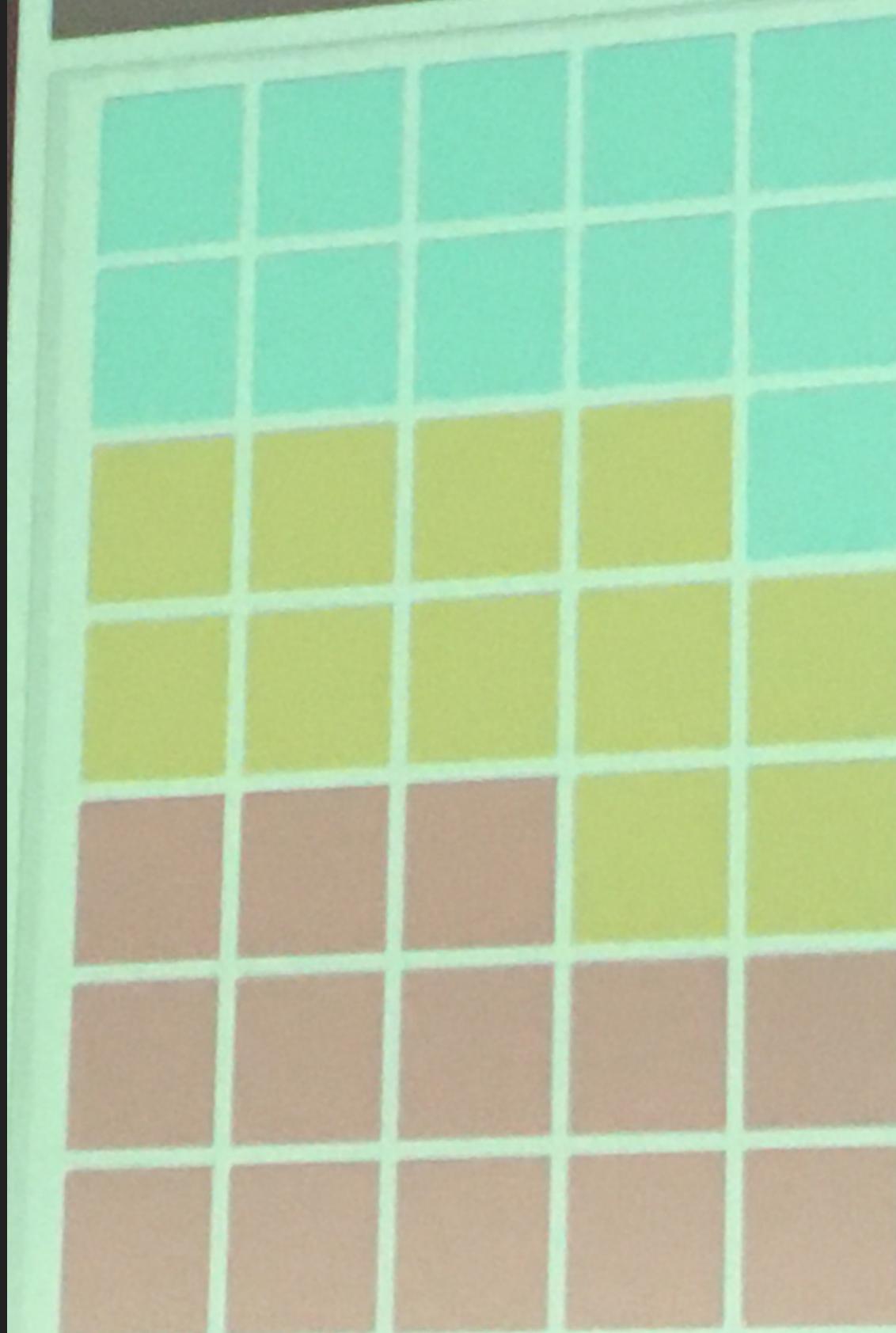
Our path to better science in less time using open data science tools

Julia S. Stewart Lowndes , Benjamin D. Best, Courtney Scarborough, Jamie C. Afflerbach, Melanie R. Frazier, Casey C. O’Hara, Ning Jiang & Benjamin S. Halpern





HOW ARE WE DOING:
RESULTS REPRODUCIBILITY



11 (31%)
Reproducible

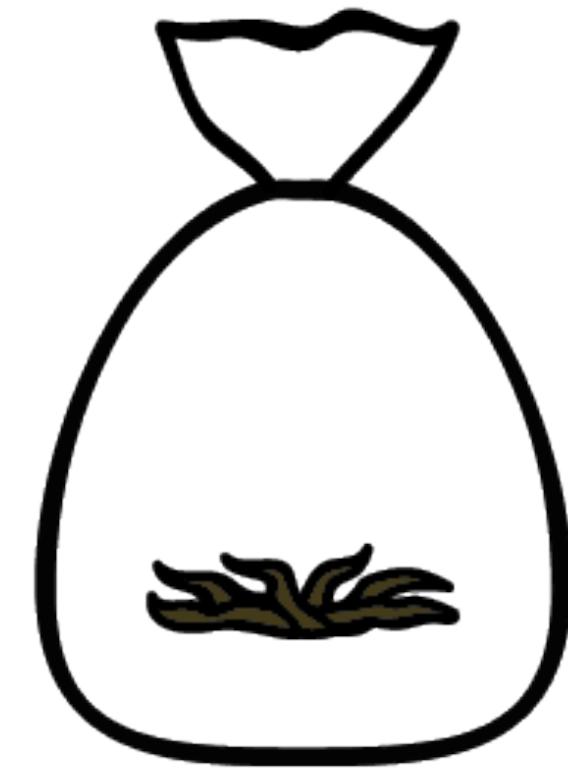
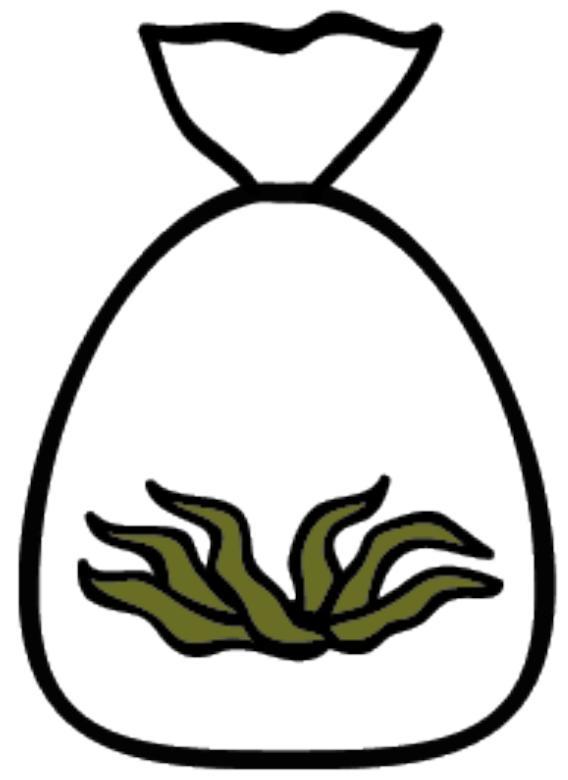
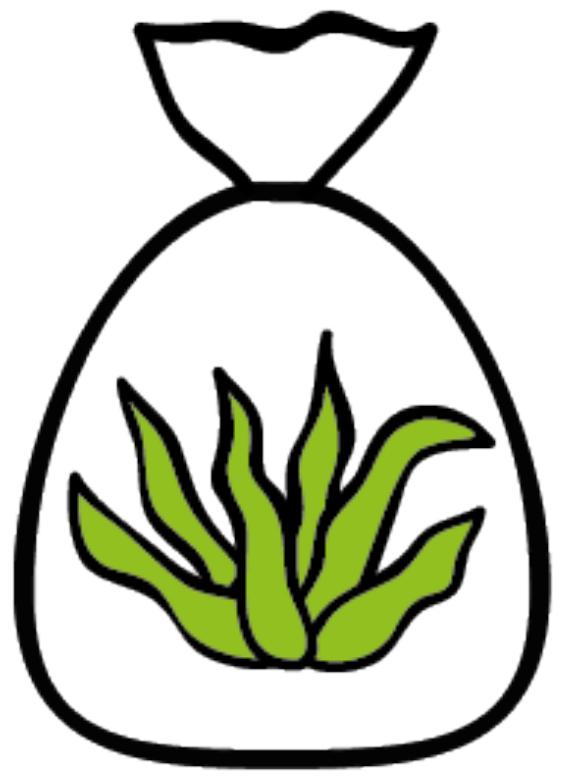
11 (31%)
**Reproducible
with author assistance**

13 (37%)
**Not fully reproducible
despite author assistance**

Hardwicke et al. (2018)

**“YOU'RE ALWAYS COLLABORATING
WITH FUTURE YOU!”**





MORE THAN A YEAR LATER... .



SIX MONTHS (AND AN R UPDATE) LATER...



#Laff

EXERCISE 2:

**DISCUSS WITH YOUR TABLE: INTRODUCE
YOURSELVES. HAVE YOU EVER TRIED TO
RERUN CODE THAT DIDN'T REPRODUCE?**

BARRIERS TO REPRODUCING CODE

1. SOFTWARE VERSIONS
2. R VERSION
3. OPERATING SYSTEMS
4. ETC.

ORGANISATION

SEPARATE DATA AND CODE

ORGANISATION

SEPARATE DATA AND CODE

DOCUMENTATION

SPECIFY DEPENDENCIES

ORGANISATION

SEPARATE DATA AND CODE

DOCUMENTATION

SPECIFY DEPENDENCIES

AUTOMATION

MASTER SCRIPT

ORGANISATION SEPARATE DATA AND CODE

DOCUMENTATION SPECIFY DEPENDENCIES

AUTOMATION MASTER SCRIPT

DISSEMINATION SHARE!

COMPENDIUM:

- ▶ ORGANISATION SYSTEM FOR THE DIFFERENT ELEMENTS THAT MAKE UP A RESEARCH PROJECT
- ▶ AS A MEANS FOR DISTRIBUTING, MANAGING AND UPDATING THE COLLECTION

EXERCISE 2.1:

FIRST UP! TYPE

USETHIS::USE`_BLANK_
_SLATE()`

AND FOLLOW THE INSTRUCTIONS.

EXERCISE 2.2:
**YOUR TURN! CREATE A REPO ON GITHUB, THEN CLONE
IT, AND START A NEW R PROJECT BASED ON THAT REPO
(FILE > NEW PROJECT > VERSION CONTROL > GIT)**

EXERCISE 2.3:

CREATE DESCRIPTION, LICENSE, AND README FILES

EXERCISE 2.4:
**YOUR TURN! CREATE A FILE STRUCTURE,
MAKE SOME FAKE DATA, START YOUR MASTER
SCRIPT**

FUNCTIONS:

WHY?

- ▶ AVOID REPETITION
- ▶ EASIER DEBUGGING
- ▶ REUSABLE

EXERCISE 2.1:
DISCUSS WITH YOUR TABLE: HOW DO YOU
WORK WITH FUNCTIONS? WHICH FUNCTIONS
DO YOU LIKE USING? WHY? WHEN DO YOU
THINK WE SHOULD WRITE FUNCTIONS?

EXERCISE 2.2:

**DOWNLOAD THE FUNCTIONS.R SCRIPT AT [HTTPS://
GITHUB.COM/ESA-QERC/ESA2019 WORKSHOP.](https://github.com/ESA-QERC/ESA2019_WORKSHOP)**

**REWORK THIS SCRIPT BY INCLUDING FUNCTIONS
WHERE YOU THINK THEY'RE APPROPRIATE.**

EXERCISE 2.3:

**DISCUSS WITH YOUR TABLE: WHICH BITS OF THE
SCRIPT DID YOU TARGET? WHY? HOW MANY
FUNCTIONS DID YOU WRITE? HOW OFTEN IS EACH
USED? WERE THERE COMMON APPROACHES?**

EXERCISE 2.4:
**REVISIT YOUR R SCRIPT FROM EARLIER (OR WORK
WITH SOME OF YOUR OWN CODE). IDENTIFY ONE
FUNCTION THAT COULD BE MORE FLEXIBLE WITH
MINOR CHANGES TO ITS ARGUMENTS OR STRUCTURE.**

VERSION CONTROL

UNTITLED 138.docx
Untitled 241.doc
Untitled 138 copy.docx
Untitled 138 copy 2.docx
Untitled 139.docx
Untitled 40 MOM ADDRESS.jpg
Untitled 242.doc
Untitled 243.doc
Untitled 243 IMPORTANT.doc
Untitled 41.fan



PROTIP: NEVER LOOK IN SOMEONE
ELSE'S DOCUMENTS FOLDER.

EXERCISE 3.1:
**DISCUSS WITH YOUR TABLE: WHAT
STRATEGIES HAVE YOU USED TO TRACK
MULTIPLE VERSIONS OF A FILE?**

EXERCISE 3.2:

**YOUR TURN! MAKE A CHANGE TO A
FILE, THEN TRY STAGING, COMMITTING
(WITH A MESSAGE).**

EXERCISE 3.3:
**CHECK OUT YOUR DIFF HISTORY, AND
BACK UP YOUR WORK TO GITHUB BY
PUSHING TO YOUR REMOTE.**

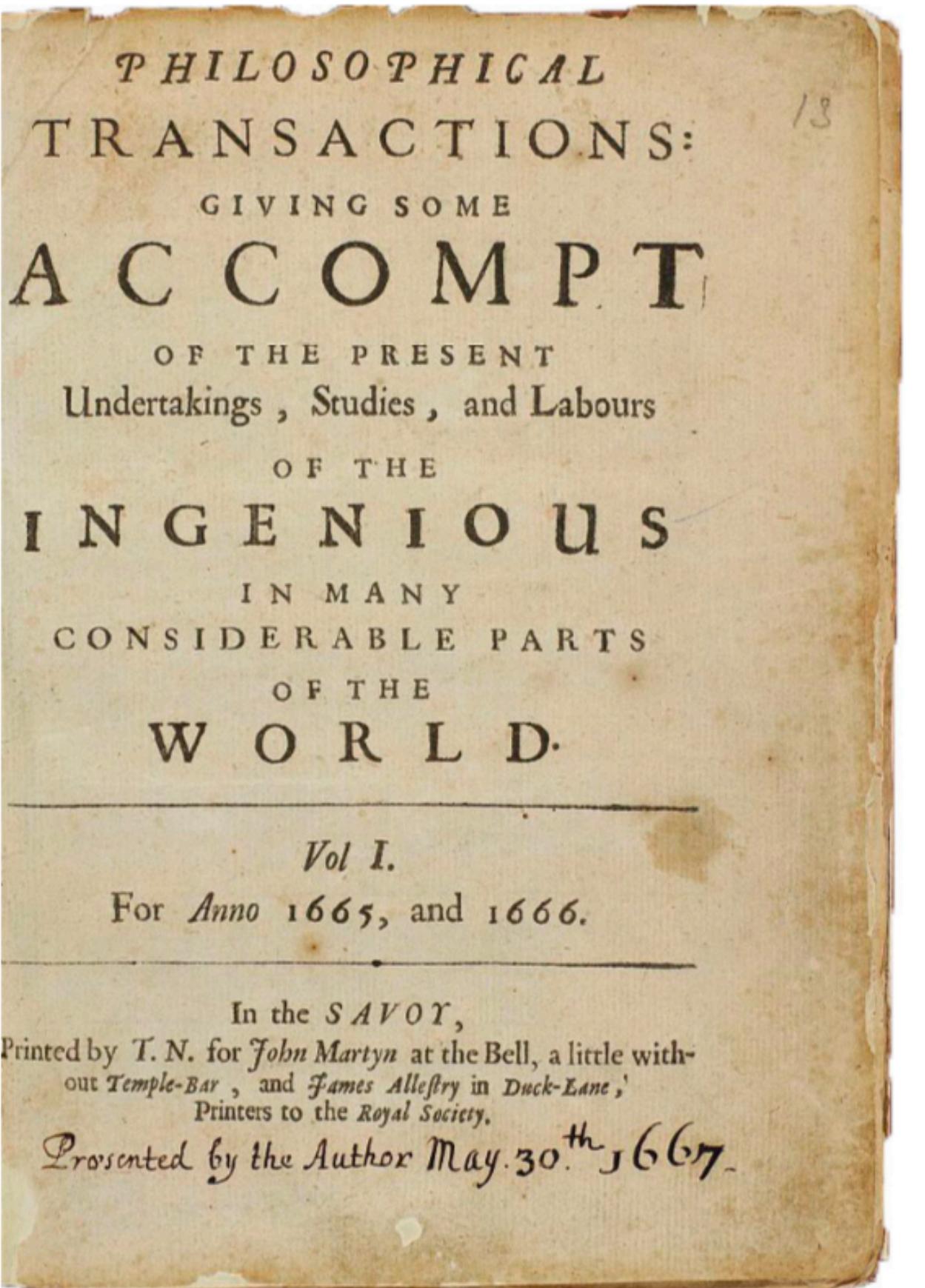
EXERCISE 3.4:

PRACTICE STAGING MULTIPLE FILES AT ONCE, AND IGNORING OTHERS

DISSEMINATING DATA & CODE

- ▶ YOUR ANALYSIS IS DONE & PAPER ACCEPTED.
NOW WHAT?

1666 – 2010



2010 →

100 YEARS Journal of Ecology



Journal of Ecology 2014, 102, 1266–1272

doi: 10.1111/1365-2745.12260

How much of the world is woody?

Richard G. FitzJohn^{1,2†}, Matthew W. Pennell^{3,4*,†}, Amy E. Zanne^{5,6}, Peter F. Stevens^{7,8}, David C. Tank⁹ and William K. Cornwell^{9,10}

¹Biodiversity Research Centre and Department of Zoology University of British Columbia, Vancouver, BC V6G 1Z4, Canada; ²Department of Biological Sciences Macquarie University, Sydney, NSW 2109, Australia; ³Department of Biological Sciences and Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA; ⁴National Evolutionary Synthesis Center, Durham, NC 27705, USA; ⁵Department of Biological Sciences, George Washington University, Washington, DC 20052, USA; ⁶Center for Conservation and Sustainable Development Missouri Botanical Garden, Louis, MO 63121, USA; ⁷Department of Biology, University of Missouri, St. Louis, MO 63166, USA; ⁸Missouri Botanical Garden, PO Box 299, St Louis, MO 63166-0299, USA; ⁹Department of Systems Ecology, VU University, 1081 HV, Amsterdam, The Netherlands; and ¹⁰Evolution & Ecology Research Centre, School of Biological Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

Summary

1. The question posed by the title of this study is a basic one, and it is surprising that the answer is not known. Recently, assembled trait data sets provide an opportunity to address this, but scaling these data sets to the global scale is challenging because of sampling bias. Although we currently know the growth form of tens of thousands of species, these data are not a random sample of global diversity; some clades are exhaustively characterized, while others we know little to nothing about.
2. Starting with a data base of woodiness for 39 313 species of vascular plants (12% of taxonomically resolved species, 59% of which were woody), we estimated the status of the remaining taxonomically resolved species by randomization. To compare the results of our method to conventional wisdom, we informally surveyed a broad community of biologists. No consensus answer to the question existed, with estimates ranging from 1% to 90% (mean: 31.7%).
3. After accounting for sampling bias, we estimated the proportion of woodiness among the world's vascular plants to be between 45% and 48%. This was much lower than a simple mean of our data set and much higher than the conventional wisdom.
4. *Synthesis.* Alongside an understanding of global taxonomic diversity (i.e. number of species globally), building a functional understanding of global diversity is an important emerging research direction. This approach represents a novel way to account for sampling bias in functional trait data sets and to answer basic questions about functional diversity at a global scale.

Key-words: data bases, determinants of plant community diversity and structure, functional diversity, herbaceousness, macroecology, sampling bias, woodiness

Introduction

The last two thousand years of research into wood since Theophrastus classified his garden have uncovered its origin in the early Devonian (~400 Mya; Gerrienne *et al.* 2011); that prevalence of woodiness varies with climate (Moles *et al.* 2009); that wood has been lost many times in diverse groups, both extant and extinct (Judd, Sanders & Donoghue 1994), often as an adaptation to freezing temperatures (Zanne *et al.* 2014); that it has also been gained many times, particularly on island systems (Carquist 1974; Givnish 1998); and that many different forms of pseudo-woody growth habit have appeared across groups that have lost true woodiness or diverged before true woodiness evolved (Cornwell *et al.* 2009). We know about its mechanical properties and developmental pathways, its patterns of decomposition and their effects on ecosystem function (Cornwell *et al.* 2009) and that

*Correspondence author. E-mail: mwpennell@gmail.com

†These authors contributed equally.

© 2014 The Authors. Journal of Ecology © 2014 British Ecological Society



data: [10.5061/dryad.v7m14.2](https://doi.org/10.5061/dryad.v7m14.2)
code: github.com/richfitz/wood
blog: ropensci.org/blog/2014/06/09/reproducibility/

Knowledge is
shared, peer re-
viewed, attributable

DISCUSS IDEAL SUPPLEMENTS TO PAPER?

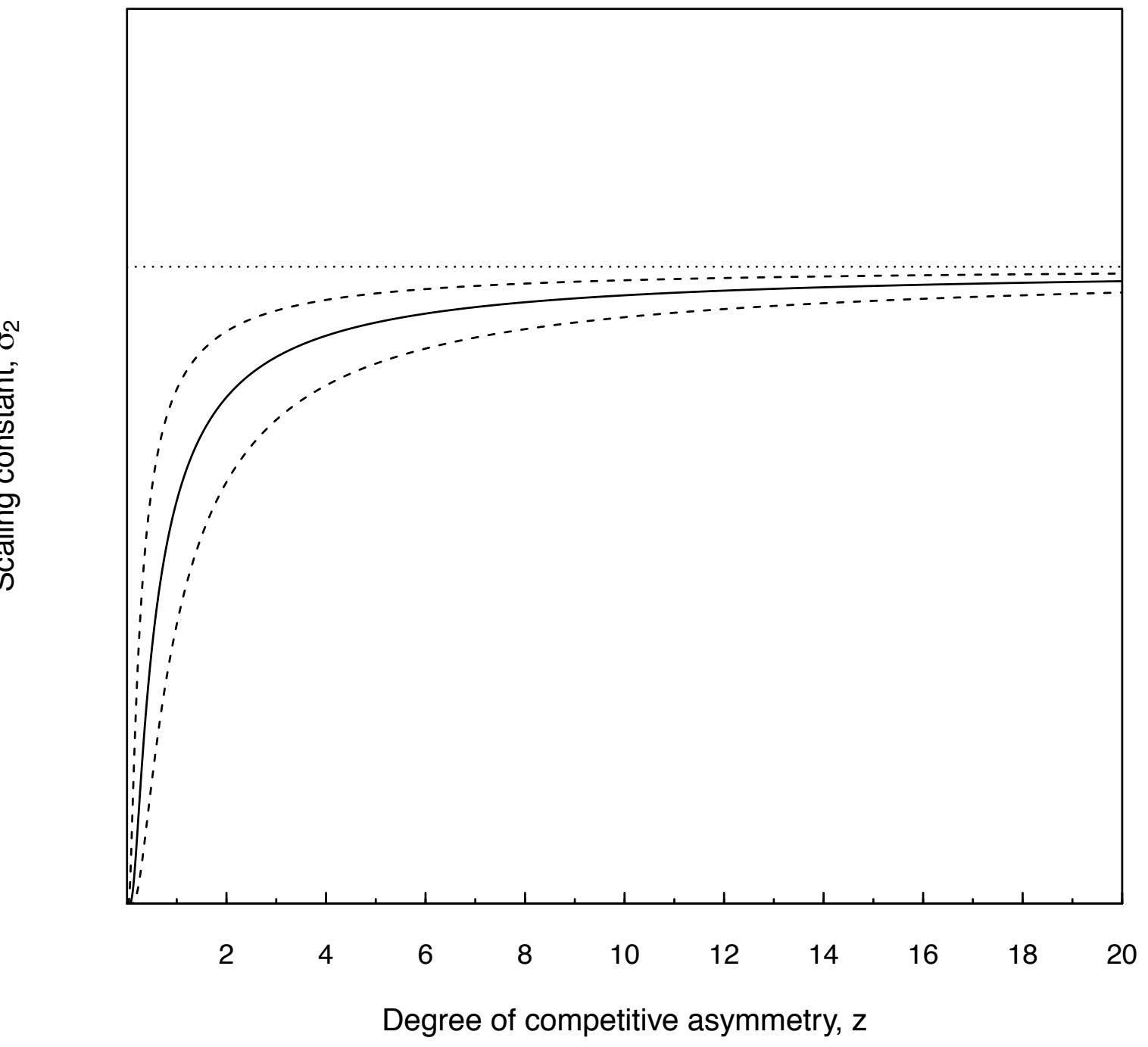
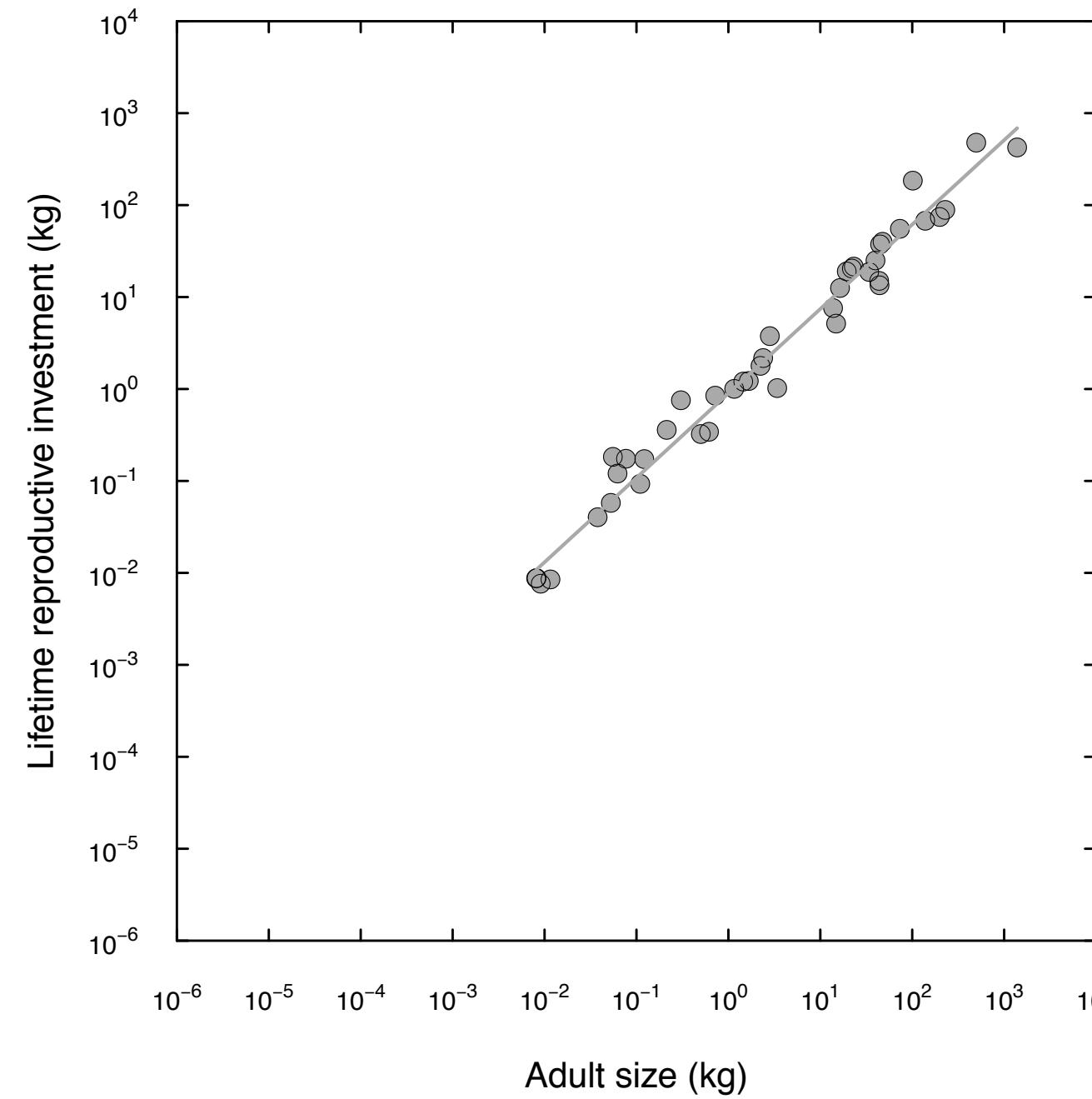
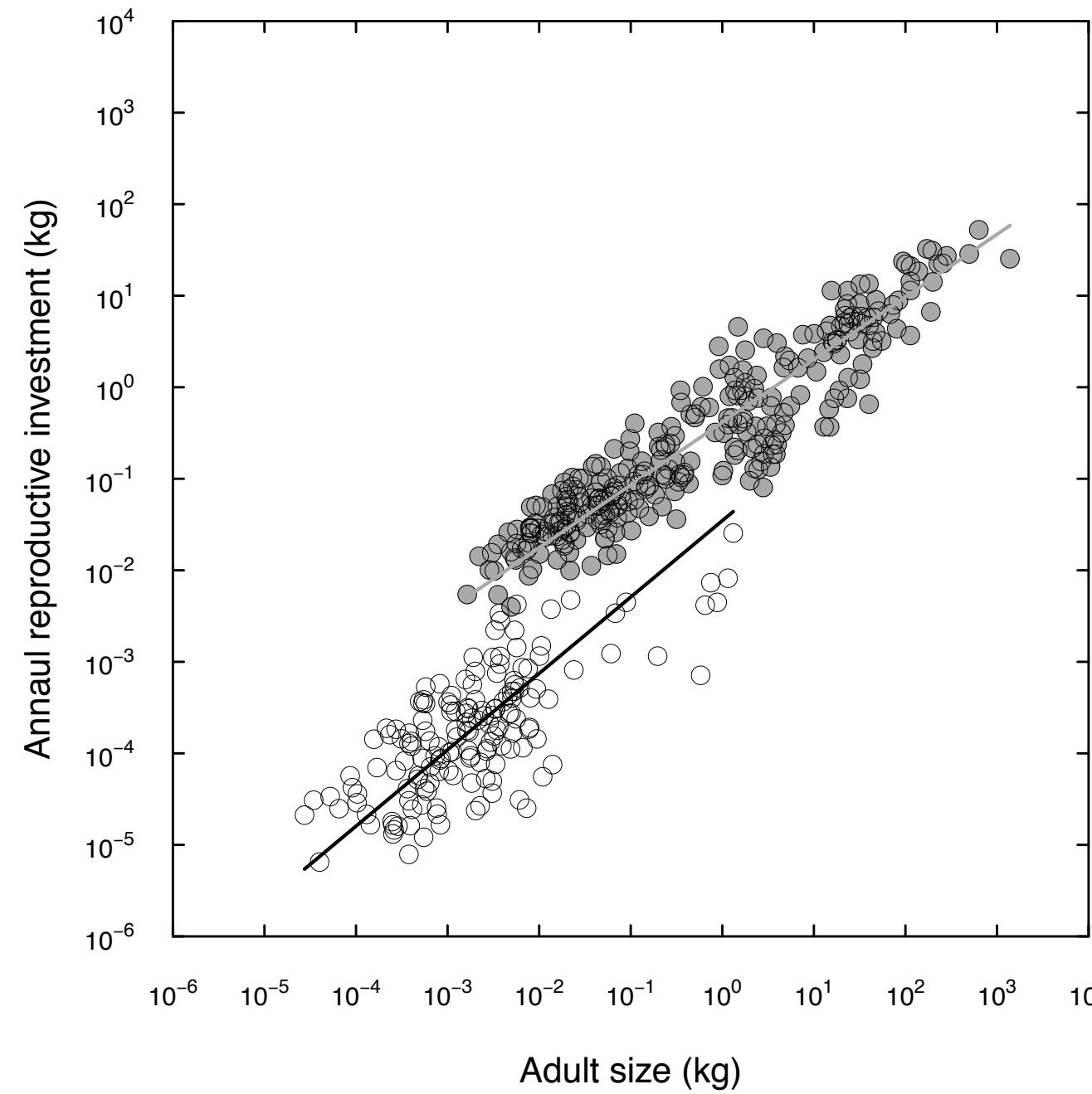
- ▶ WHAT WOULD YOU LIKE TO SEE?

MY IDEAL PAPER

- ▶ All analyses and data in git repository, on GitHub
 - ▶ Code runs, produces figures, tables
 - ▶ Nice readme
 - ▶ License
- ▶ Publication version of code and data archived in long-term repository
- ▶ Links in paper
- ▶ Anyone (including author) can reproduce & build on research

EXERCISE 4.1: REPRODUCE RESULTS OF FALSTER ET AL 2008

https://github.com/dfalster/Falster_2008_AmNat_offspring_model/



EXERCISE 4.2:

LOG INTO ZENODO AND ‘WATCH’ YOUR PROJECT REPO

The screenshot shows the Zenodo GitHub integration interface. At the top, there's a blue header bar with the Zenodo logo, a search bar, an upload button, and a communities link. On the right, there's a user profile dropdown showing 'daniel.falster@unsw.edu.au'. Below the header, the URL 'Home / Account / GitHub' is visible. On the left, a sidebar titled 'Settings' lists options: Profile, Change password, Security, Linked accounts, Applications, Shared links, and GitHub (which is highlighted in blue). The main content area is titled 'GitHub Repositories' and shows a 'Get started' section with three steps: 1. Flip the switch (with a note about selecting a repository and turning on automatic preservation), 2. Create a release (with a note about going to GitHub to create a release and Zenodo automatically downloading it), and 3. Get the badge (with a note about Zenodo registering a DOI for each new release). A 'Sync now ...' button is also present.

zenodo

Search Upload Communities

daniel.falster@unsw.edu.au

Home / Account / GitHub

Settings

- Profile
- Change password
- Security
- Linked accounts
- Applications
- Shared links
- GitHub**

GitHub Repositories (updated now) Sync now ...

Get started

- 1 Flip the switch**

Select the repository you want to preserve, and toggle the switch below to turn on automatic preservation of your software.

ON
- 2 Create a release**

Go to GitHub and [create a release](#). Zenodo will automatically download a .zip-ball of each new release and register a DOI.
- 3 Get the badge**

After your first release, a DOI badge that you can include in GitHub README will appear next to your repository below.

DOI **10.5281/zenodo.8475**
(example)

CREATING A GITHUB RELEASE

- ▶ A 'RELEASE' = A SNAPSHOT OF REPO
- ▶ CODE AS ZIP FILE
- ▶ CAN ALSO INCLUDE PRODUCTS PRODUCED BY CODE
(THINK SOFTWARE BINARIES)

- ▶ EXERCISE 4.3 — CREATE A RELEASE OF YOUR REPO

EXERCISE 4.4:

ARCHIVE A SNAPSHOT YOUR REPO ON ZENODO (DONE)

GET THE DOI AND ADD TO YOUR GITHUB README!

CELEBRATE!

LINKS IN YOUR PAPER

Received: 11 September 2017 | Accepted: 24 January 2018

DOI: 10.1111/1365-2745.12974

Check for
updates

LINKING ORGANISMAL FUNCTIONS, LIFE HISTORY STRATEGIES AND POPULATION PERFORMANCE

Journal of Ecology



Investment in reproduction for 14 iteroparous perennials is large and associated with other life-history and functional traits

Elizabeth Hedi Wenk¹ | Konrad Abramowicz² | Mark Westoby¹ | Daniel S. Falster³

AUTHORS' CONTRIBUTIONS

D.S.F. and E.H.W. conceived the idea; E.H.W. collected the data; all authors analysed the data; E.H.W. and D.S.F. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Code and data deposited in Zenodo: <https://doi.org/10.5281/zenodo.1183416> (Wenk, Abramowicz, Westoby, & Falster, 2018).

ORCID

Daniel S. Falster <http://orcid.org/0000-0002-9814-092X>

- Franco, M., & Silvertown, J. (2004). A comparative demography of plants based upon elasticities of vital rates. *Ecology*, 85, 531–538. <https://doi.org/10.1890/02-0651>
- Gaillard, J.-M., Yoccoz, N. G., Lebreton, J.-D., Bonenfant, C., Devillard, S., & Loison, A., ... Allaine, D. (2005). Generation time: A reliable metric to measure life-history variation among mammalian populations. *The American Naturalist*, 166, 119–123. <https://doi.org/10.1086/430330>
- Gibert, A., Gray, E. F., Westoby, M., Wright, I. J., & Falster, D. S. (2016). On the link between functional traits and growth rate: Meta-analysis shows effects change with plant size, as predicted. *Journal of Ecology*, 104, 1488–1503. <https://doi.org/10.1111/1365-2745.12594>
- Givnish, T. (1995). Plant stems: Biomechanical adaptation for energy capture and influence on species distributions. *Plant stems: Physiology and functional morphology* (pp. 3–49). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012276460-8/50003-5>
- Harper, J. L., & Ogden, J. (1970). The reproductive strategy of higher plants: I. the concept of strategy with special reference to *Senecio vulgaris*

Journal of Ecology

On the link between functional traits and growth rate: Meta-analysis shows effects change with plant size, as predicted

Anaïs Gibert*, Emma F. Gray, Mark Westoby, Ian J. Wright and Daniel S. Falster

Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia

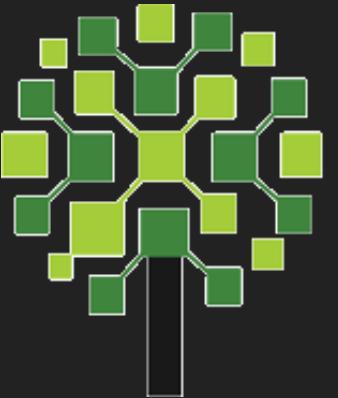
Data accessibility

The code and data for producing all figures and results in this paper are available at github.com/AnaisGibert/Growth_trait_metaanalysis. Data are available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.701q8> (Gibert *et al.* 2016).

References

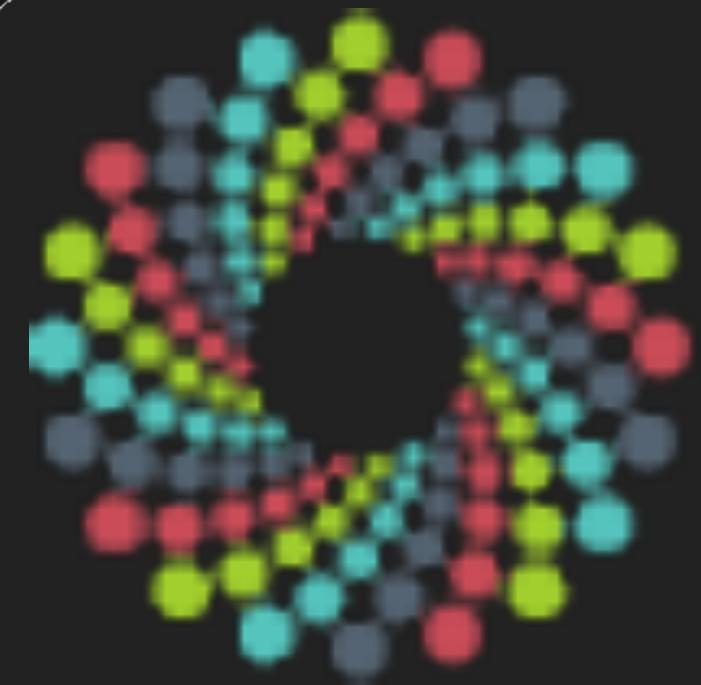
- Aiba, S.I. & Kohyama, T. (1997) Crown architecture and life-history traits of 14 tree species in a warm temperate rain forest: significance of spatial heterogeneity. *Journal of Ecology*, 85, 531–538. <https://doi.org/10.1111/j.1365-2745.1997.tb01920.x>

SOME REPOSITORIES

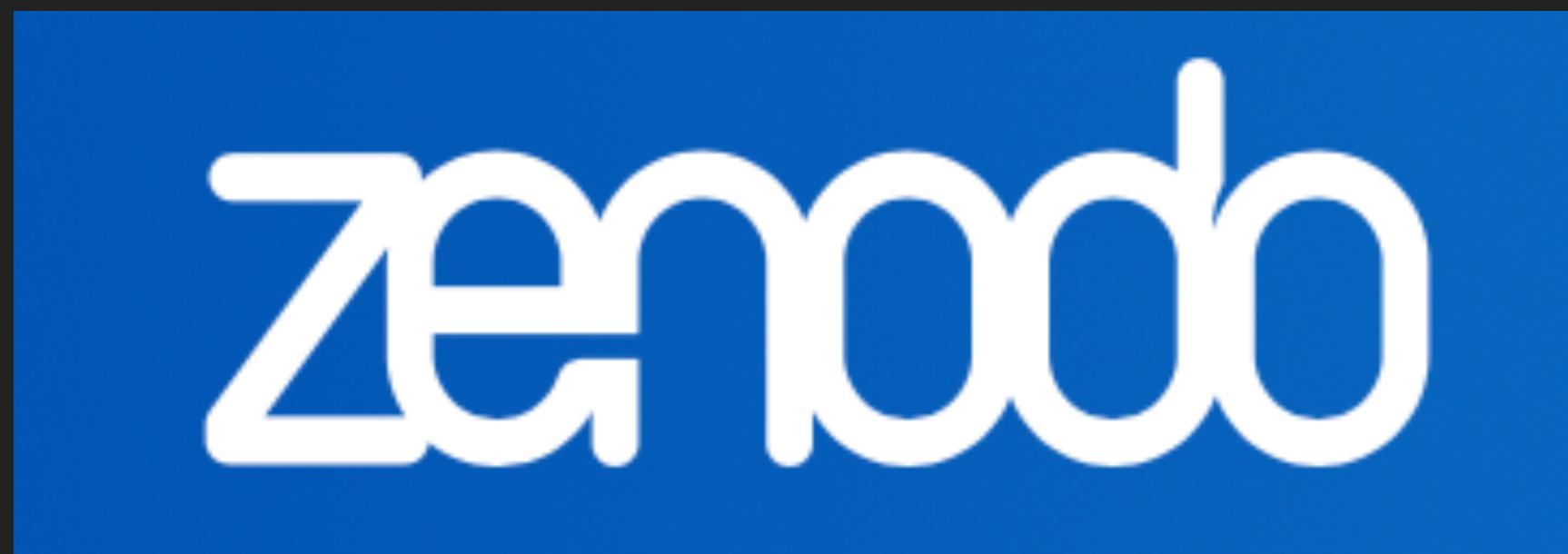


DRYAD

try-db.org



figshare



journal supplementary

MY WORKFLOW

- ▶ All analyses in git on GitHub (private) from outset
- ▶ Collaborate via GitHub
- ▶ When submitting, make public, make a release, put GitHub link in paper
- ▶ Revise, resubmit.....revise resubmit (new releases?)
- ▶ At proof stage: Archive in Zenodo
- ▶ Links in paper, DOI badge on GitHub readme

**BUT STILL, CODE CAN
BREAK!**

CONTAINERISATION TO THE RESCUE!

1. A TYPE OF VIRTUALISATION
2. MORE LIGHTWEIGHT THAN A VIRTUAL MACHINE
3. POPULARLY IMPLEMENTED WITH DOCKER

WHEN TO USE DOCKER/ CONTAINERISATION

- 1. WHEN ACTIVELY WORKING ON A
PROJECT - TO MAINTAIN PACKAGE AND
SOFTWARE VERSIONS**



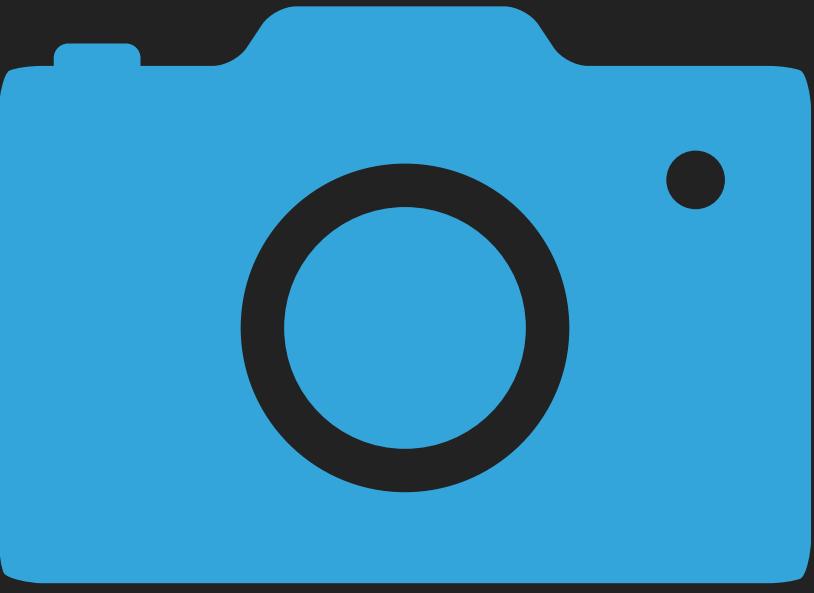
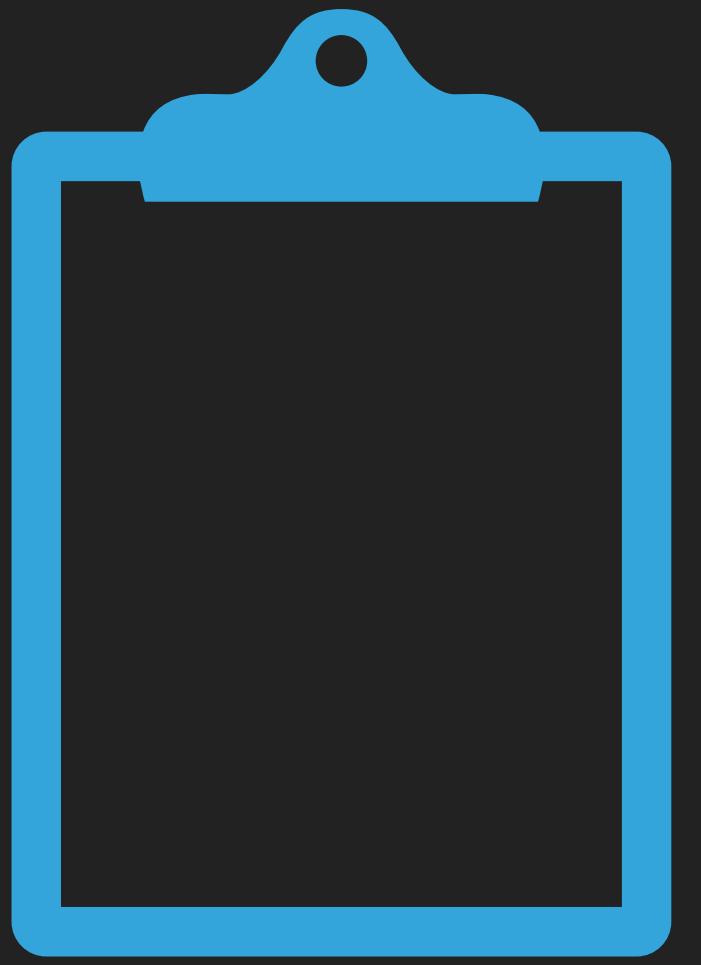
Mark Greenaway
@certifiedwaif

▼

Replying to [@certifiedwaif](#) and [@smwindecker](#)

I'll put it even more bluntly. I remember what it was like before containers came along, and I have no desire to go back.

Yes, working with containers can be complicated. But it's less complicated than trying to reckon with the constantly changing state of every person's laptop.

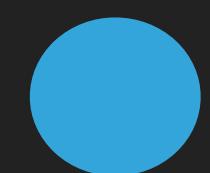






**DOCKERHUB - CONTAINS DOCKERFILES
FOR REFERENCE, AND TO BUILD UPON**

FLEXIBILITY



DOCKERFILE FROM SCRATCH

EASE

FLEXIBILITY

- DOCKERFILE FROM SCRATCH
- CONTAINERIT PACKAGE
- HOLEPUNCH PACKAGE

EASE

HOLEPUNCH R PACKAGE BY KARTHIK RAM

HOLEPUNCH HAS TWO MAIN ADVANTAGES:

1. AUTOMATICALLY PREPARES A DOCKER FILE USING USING METADATA CONTAINED IN A DESCRIPTION FILE

HOLEPUNCH HAS TWO MAIN ADVANTAGES:

1. AUTOMATICALLY PREPARES A DOCKER FILE USING USING METADATA CONTAINED IN A DESCRIPTION FILE
2. LAUNCHES CONTAINER WITHOUT DOWNLOADING DOCKER BY LEVERAGING BINDER

WHEN TO USE DOCKER/ CONTAINERISATION

- 1. WHEN ACTIVELY WORKING ON A
PROJECT - TO MAINTAIN PACKAGE AND
SOFTWARE VERSIONS**
- 2. TO SHARE YOUR PROJECT AND ENABLE
OTHERS TO EASILY REPRODUCE IT**

ULTIMATELY, REPRODUCIBLE CODE HAS:
FEWER MISTAKES & LESS DUPLICATION

ULTIMATELY, REPRODUCIBLE CODE HAS:
FEWER MISTAKES & LESS DUPLICATION

IS:
MORE INTERPRETABLE & EASILY SHARED

ULTIMATELY, REPRODUCIBLE CODE HAS:
FEWER MISTAKES & LESS DUPLICATION

IS:
MORE INTERPRETABLE & EASILY SHARED

AND SUPPORTS:
MORE RELIABLE AND ROBUST SCIENCE

THANK YOU!

RESOURCES AND SLIDES AT:

[GITHUB.COM/ESA-QERC/ESA2019 WORKSHOP](https://github.com/ESA-QERC/ESA2019_WORKSHOP)

NICK.GOLDING.RESEARCH@GMAIL.COM
 @_NICKGOLDING_

DANIEL.FALSTER@UNSW.EDU.AU
 @ADAPTIVE_PLANT

JIAN.YEN@UNIMELB.EDU.AU

SM.WINDECKER@UNIMELB.EDU.AU
 @SMWINDECKER

