SARAS WINDECKER @SMWINDECKER

GOLD STAR REPRODUCIBILITY: STRAIGHTFORWARD CONTAINERISATION

NULLIUS IN VERBA

TAKE NOBODY'S WORD FOR IT.

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

1,500 scientists lift the lid on reproducibility

Survey sheds light on the 'crisis' rocking research.

Monya Baker

25 May 2016

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

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**
2. **QUESTIONABLE RESEARCH PRACTICES IN 'PUBLISH OR PERISH' CULTURE**

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

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

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

Corrected: 28 July 2016

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

The sc
what c
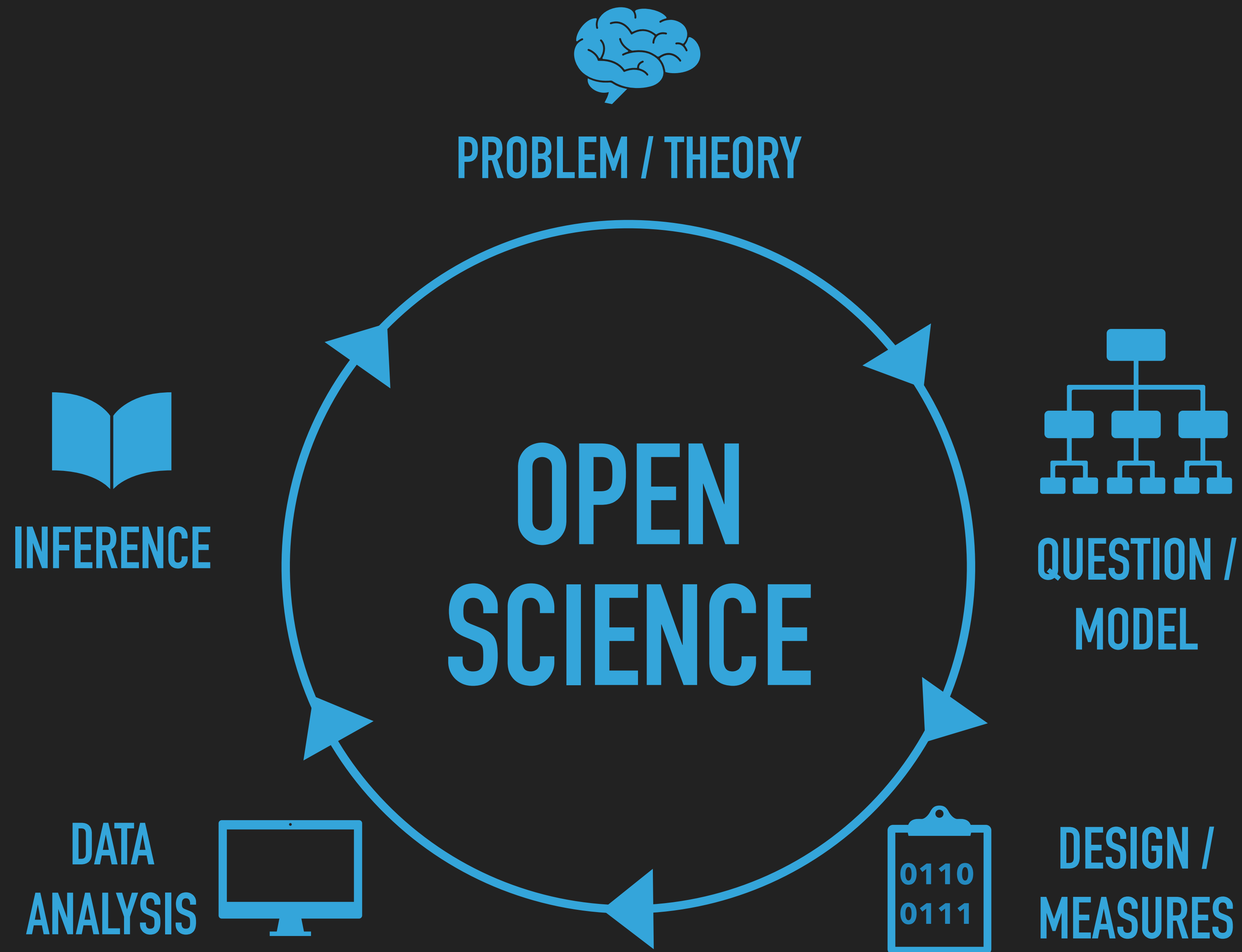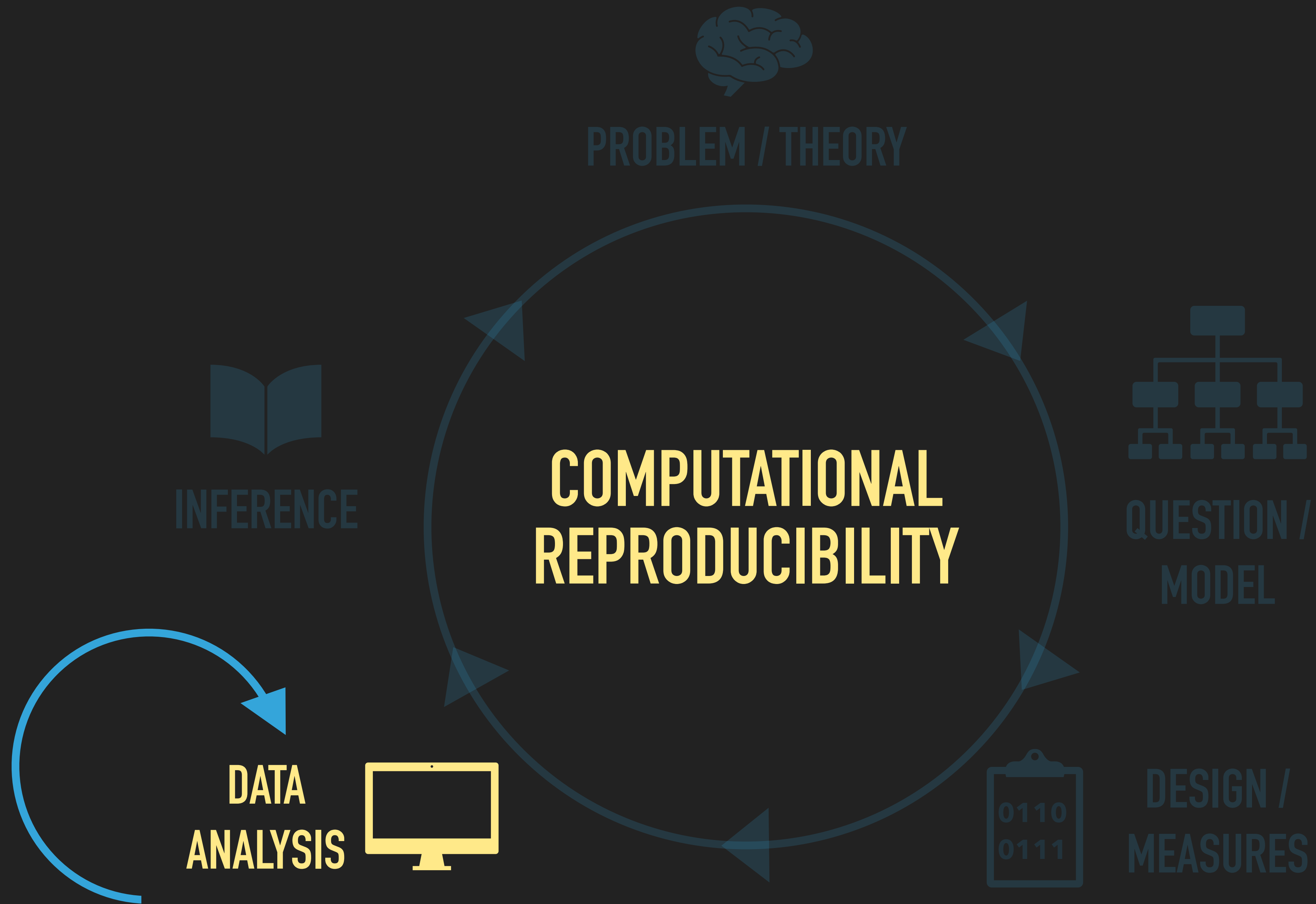
March 15, 2017 8.49p

nature
ecology & evolution

Perspective | Published: 23 May 2017

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

PROBLEM / THEORY

INFERENCE

COMPUTATIONAL
REPRODUCIBILITY

QUESTION /
MODEL

DATA
ANALYSIS

DESIGN /
MEASURES

0110
0111

Adapted from Rachel Searston @SearstonR

COMPENDIUM:

▸ ORGANISATION SYSTEM FOR THE DIFFERENT ELEMENTS THAT MAKE UP A RESEARCH PROJECT

▸ AS A MEANS FOR DISTRIBUTING, MANAGING AND UPDATING THE COLLECTION

Gentleman, R. and Temple Lang, D. (2004)

ORGANISATION   SEPARATE DATA AND CODE

DOCUMENTATION   SPECIFY DEPENDENCIES

AUTOMATION   MASTER SCRIPT

DISSEMINATION   SHARE!

Thanks to April Clyburne-Sherin at CodeOcean

MORE THAN A YEAR LATER...

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

#Laff

WHY WOULDN'T WELL-WRITTEN, WELL-ORGANISED CODE REPRODUCE?

BARRIERS TO REPRODUCING CODE

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

ORGANISATION SEPARATE DATA AND CODE

DOCUMENTATION SPECIFY DEPENDENCIES

AUTOMATION MASTER SCRIPT

DISSEMINATION SHARE!

Thanks to April Clyburne-Sherin at CodeOcean

CONTAINERISATION TO THE RESCUE!

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

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

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.
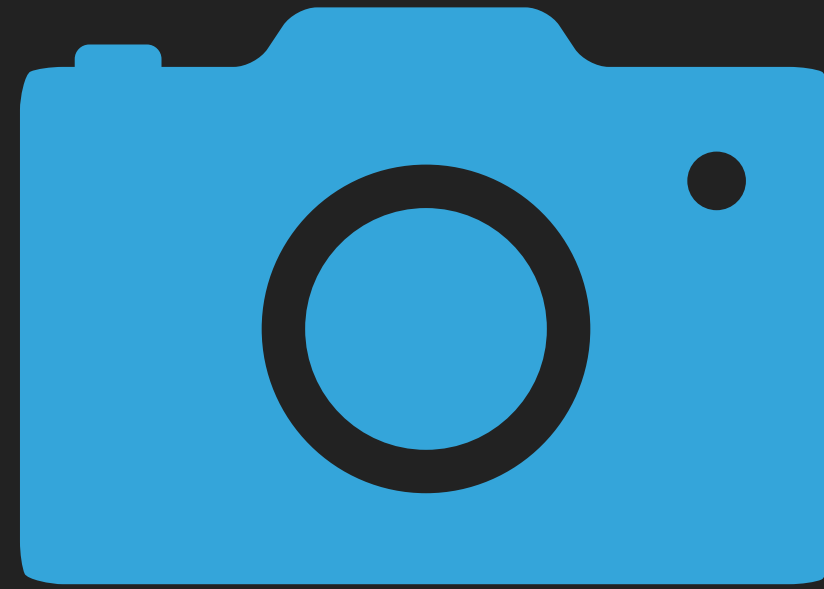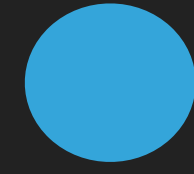
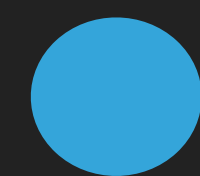# DOCKERHUB – CONTAINS DOCKERFILES FOR REFERENCE, AND TO BUILD UPON

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

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

THE FIRST PERSON THAT WILL RERUN YOUR ANALYSIS IS YOU.

THANK YOU!

RESOURCES AND SLIDES AT:

GITHUB.COM/SMWINDECKER/HOLEPUNCH_RLADIES

KARTHIK RAM FOR HOLEPUNCH
RLADIES MELBOURNE

SM.WINDECKER@UNIMELB.EDU.AU
@SMWINDECKER