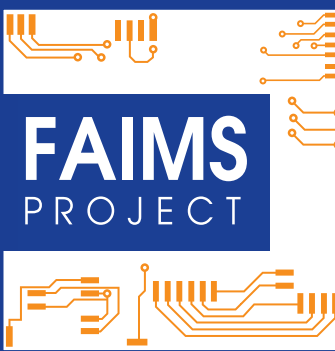


Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

By: Dr Brian Ballsun-Stanton

(Twitter: @DenubisX)

Technical Director

Field Acquired Information Management Systems
Project

Department of Ancient History
Macquarie University

This presentation taken from notes on a forthcoming chapter called: *Open Data and the Danger of Sympathetic Magic* by: Dr Brian Ballsun-Stanton, Macquarie University, Technical Director of the FAIMS project and Georgia Burnett, Macquarie University. Typeset with ConT_EXt.

Presentation source available on github.com/FAIMS/OpenDataPresentation.

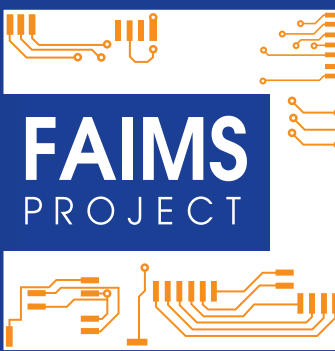


Australian Government
Australian Research Council
LIEF AWARD: LE140100151



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Collecting data is hard, reusing it is harder

- I help people deal with data collection while offline
- Collecting data in the middle of nowhere is a huge pain
- Even bigger challenge: working with already collected data



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY • AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO

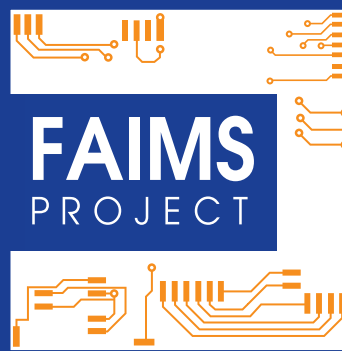


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Show of hands:

How many of you have tried to use someone else's data?



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY · AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO

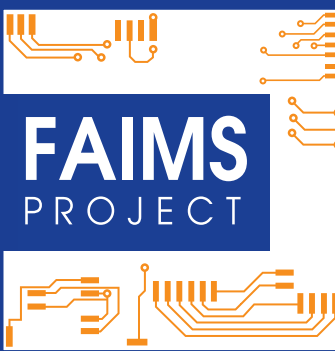


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Current Challenges of Open Data

- No rewards for quality open data releases
- What's a "knowledge-creation framework?"
- Write-once Read-once
- No Context



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY · AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar



OCHRE Data Service
of the Oriental Institute
tDAR

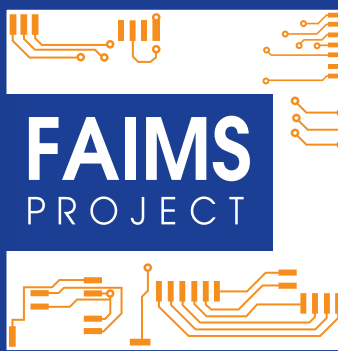


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Standing on the Shoulders of Giants

... should include standing on their data.

- Valuable data is valuable to all
- If only arguments of value, so much is lost



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY • AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO

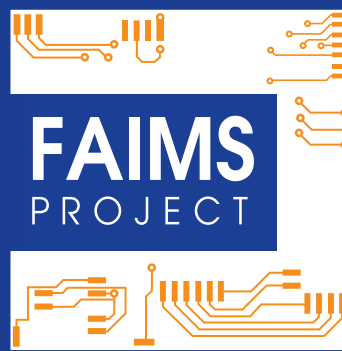


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Sympathetic Magic

"PERHAPS the most familiar application of the *principle that like produces like* is the attempt which has been made by many peoples in many ages to injure or destroy an enemy by injuring or destroying an image of him, in the belief that, just *as the image suffers, so does the man*, and that when it perishes he must die. "

Sir J.G. Frazier, The Golden Bough (1922)



Australian Government

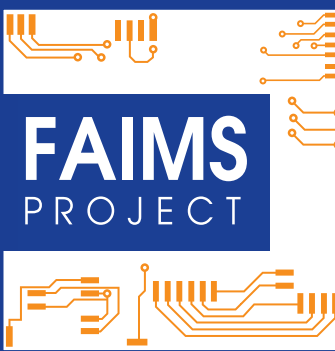
Australian Research Council

LIEF AWARD: LE140100151



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

What?

- Data is the product of methodology and analysis
- Methodology of that which to be recorded
- Methodology of how to record observations
- Methodology of analysis
- Methodology of the mechanics of analysis



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY · AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO

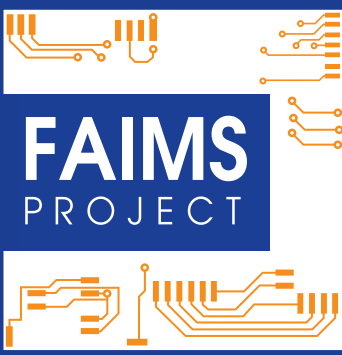


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Mere presentation of data
is sympathetic magic:

The manipulation of data without tests, context and mechanisms is
the manipulation of the data's sympathetic image.



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY • AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY OF
QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO

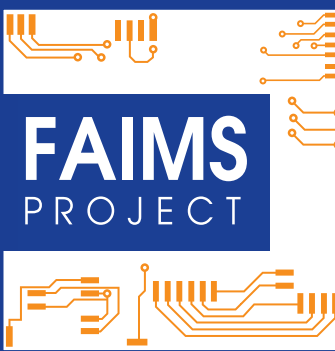


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Data as a problem of sympathetic magic

- We supply the image of the data, expecting it to be useful
- The true utility is the structure of the data and its methods of generation, storage, and analysis, not just a magic charm CSV file



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY · AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO

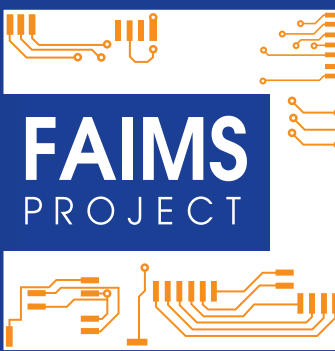


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Goals of open data

- Complete datasets online in curated spaces
 - Includes metadata for more accurate reuse
 - Curation prevents data format and structure from becoming obsolete or deteriorating by publishing data alongside scholarship
- Data that is free of cost
- Data that is free of copyright/contractual obligation



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY • AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO

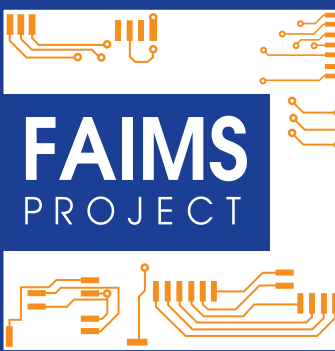


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Tests for quality

- Test: Is a re-user able to add to a dataset without outside consultation?
- Test: can the re-user re-run initial data analysis of authors?



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY · AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY OF
QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar



OCHRE Data Service
of the Oriental Institute
tDAR

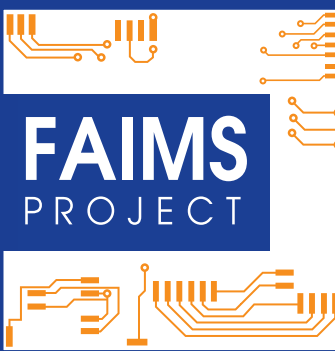


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Goals of open, useful, data

- What does the data mean to those who recorded it?
 - Documentation of mechanisms of recording and analysis
 - Documentation of judgement calls and design decisions
- Addresses epistemological questions
- Does not require out-of-band knowledge
- Does not require knowledge of conclusions before run



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY • AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar



OCHRE Data Service
of the Oriental Institute
tDAR

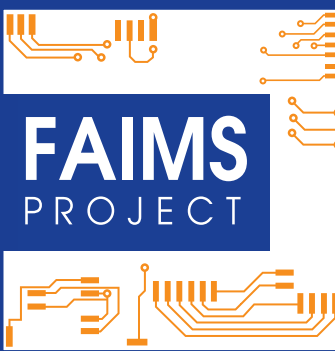


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Ethical imperatives

- Publishing data paid for by taxpayers for the "public good"
- Reasons not to hoard data
- Long-term, sustainable data preservation



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY · AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar



OCHRE Data Service
of the Oriental Institute
tDAR

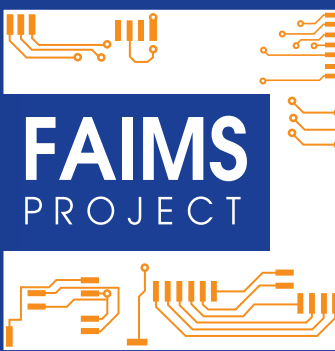


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Bigger Picture

- Change in publishing climate
- This publishing-of-analytic-programs work is adjunct to a paper
- We do this because we want other people to steal our research methods for the good of knowledge
- Ideas are cheap, execution is hard
- Research is very expensive



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY • AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar



OCHRE Data Service
of the Oriental Institute
tDAR

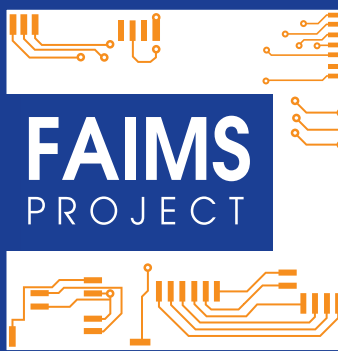


ads
ARCHAEOLOGY
DATA SERVICE



Open Data, The Reproducibility Problem, and Sympathetic Magic

Brian Ballsun-Stanton (Department of Ancient History)



www.faims.edu.au

Conclusion

- Reusing data to answer novel questions without asking the original researchers for help (or a thumb drive) is very hard.
- Publishing only a superficial image of data, however, is sympathetic magic.



Australian Government

Australian Research Council

LIEF AWARD: LE140100151



MACQUARIE
University
SYDNEY · AUSTRALIA



UNSW
AUSTRALIA



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA



THE UNIVERSITY OF
SYDNEY



LA TROBE
UNIVERSITY



Flinders
UNIVERSITY



Southern Cross
University



nectar
INTERSECT



OCHRE Data Service
of the Oriental Institute
THE UNIVERSITY OF
CHICAGO



ads
ARCHAEOLOGY
DATA SERVICE

Presentation source available on github.com/FAIMS/OpenDataPresentation.

