



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

Software and Workflow for Reproducible Research Discussion

Garret Christensen¹

¹UC Berkeley: Berkely Initiative for Transparency in the Social Sciences
Berkeley Institute for Data Science

APHRC, Summer 2015



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

bitss.org



Outline

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

1 Introduction

2 Problems

■ Irreproducible Workflow

3 Solutions

■ Workflow

4 Conclusion



BITSS

BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Reproducibility & Transparency

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- What are problems associated with reproducibility?
- What are solutions to these problems?
- What are practical tools to implement these solutions?



Problems

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- Publication bias (see previous talk)
- Specification Searching (see previous talk)
- Data not available
- Code not available/unintelligible



Irreproducible Workflow

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- Even with the original authors' help, you can't get the data to reproduce the published results. Or you just can't find the data to begin with.
- *Journal of Money, Credit, and Banking* Project. (Dewald et al., AER 1986)
- Martin Feldstein on Social Security and private savings, Reinhart and Rogoff on debt and GDP growth.



Solutions

BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- Study Registry (see previous talk)
- Pre-Analysis Plan (see previous talk)
- Reproducible Workflow
 - Literate Programming
 - Data Sharing



Solutions

BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- Study Registry (see previous talk)
- Pre-Analysis Plan (see previous talk)
- Reproducible Workflow
 - Literate Programming
 - Data Sharing



Solutions

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- Study Registry (see previous talk)
- Pre-Analysis Plan (see previous talk)
- Reproducible Workflow
 - Literate Programing
 - Data Sharing



Solutions

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- Study Registry (see previous talk)
- Pre-Analysis Plan (see previous talk)
- Reproducible Workflow
 - Literate Programming
 - Data Sharing



Solutions

Short Paper
Title

Christensen

Introduction

Problems

Irreproducible
Workflow

Solutions

Workflow

Conclusion

- Study Registry (see previous talk)
- Pre-Analysis Plan (see previous talk)
- Reproducible Workflow
 - Literate Programming
 - Data Sharing



Reproducible Workflow

Short Paper
Title

Christensen

Introduction

Problems

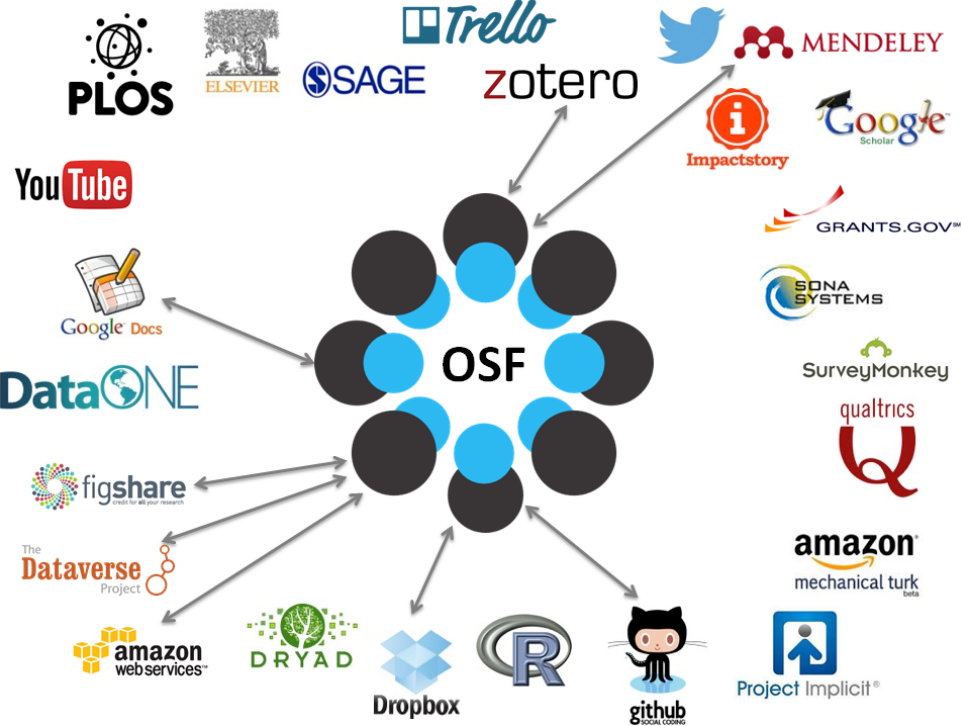
Irreproducible
Workflow

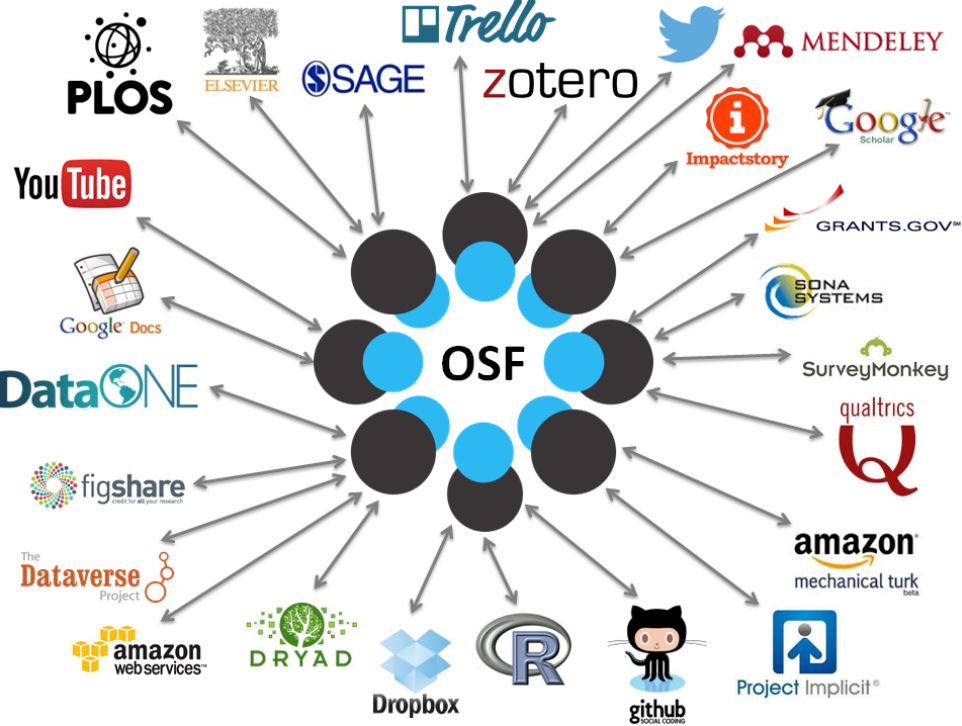
Solutions

Workflow

Conclusion

- R Markdown and R Studio to write dynamic documents.
- Version control with Github or OSF.
- Literate Programing
 - Writing code to be read by a human instead of a machine. (Loosely: commenting the hell out of it.)
- Data Sharing
 - Harvard's Dataverse





Simple tools exist to help you transparently and reproducibly take your research from beginning to end.

- Open Science Framework
- Trial Registries
- Version Control
- Dynamic Documents
- Trusted Public Data Archive

Read more in my *Manual of Best Practices in Transparent Social Science Research* on GitHub.