OPEN DATA – REPLICATION FILE

A conversation with evaluation funders

Benjamin Wood

Integra

INTRODUCTION

- Open data is not just about data
  - Replication file (Martel 2016) describes the data, statistical code, and documentation required to meaningfully use "open data"
  - Registration, data management/repository, open access

- Time to move beyond policies
  - Lack of enforcement
  - Little research community interest

- Funders hold carrots and sticks
  - Push button replication of draft results
  - Open data challenges

PANELISTS

Name	Org name	Org type	Data policy	How is your organization contributing to the open evaluation data movement?
Neeta Goel	3ie	NGO	Yes	3ie mandates pre-registration and open access to data and pubs. Replicates results before releasing final payments.
Lorena Quintero	Gov of Colombia (GoC)	Gov	Yes	The GoC publishes data from every eval, following internal protocols of verification, replication, and anonymization.
Romina Cavatassi	IFAD	IO	No	IFAD has set up a task force to discuss data policy as the intention is to move towards data sharing and transparency.
Sudharsanam Balasubramaniam	CIFF	Funder	Yes	CIFF has an open data policy. All eval data are freely accessible for re-analysis.

Why do you think open data is important for evaluation?

PANELISTS

Name	Org name	Org type	Data policy	How is your organization contributing to the open evaluation data movement?
Neeta Goel	3ie	NGO	Yes	3ie mandates pre-registration and open access to data and pubs. Replicates results before releasing final payments.
Lorena Quintero	Gov of Colombia (GoC)	Gov	Yes	The GoC publishes data from every eval, following internal protocols of verification, replication, and anonymization.
Romina Cavatassi	IFAD	IO	No	IFAD has set up a task force to discuss data policy as the intention is to move towards data sharing and transparency.
Sudharsanam Balasubramaniam	CIFF	Funder	Yes	CIFF has an open data policy. All eval data are freely accessible for re-analysis.

PROVISION OF REPLICATION FILES
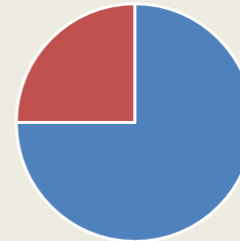

AIDS & Behavior — ▪Yes ▪No

AEJ:AE — ▪Yes ▪No
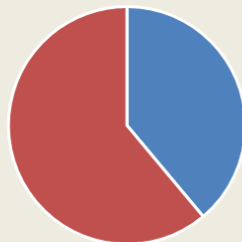
BMJ — ▪Yes ▪No

EDCC — ▪Yes ▪No

JDE — ▪Yes ▪No

JDeff — ▪Yes ▪No
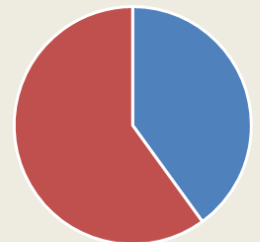
PLOS — ▪Yes ▪No

Lancet — ▪Yes ▪No

TMIH — ▪Yes ▪No

WD — ▪Yes ▪No

PRIMARY RESULTS: ALL JOURNALS
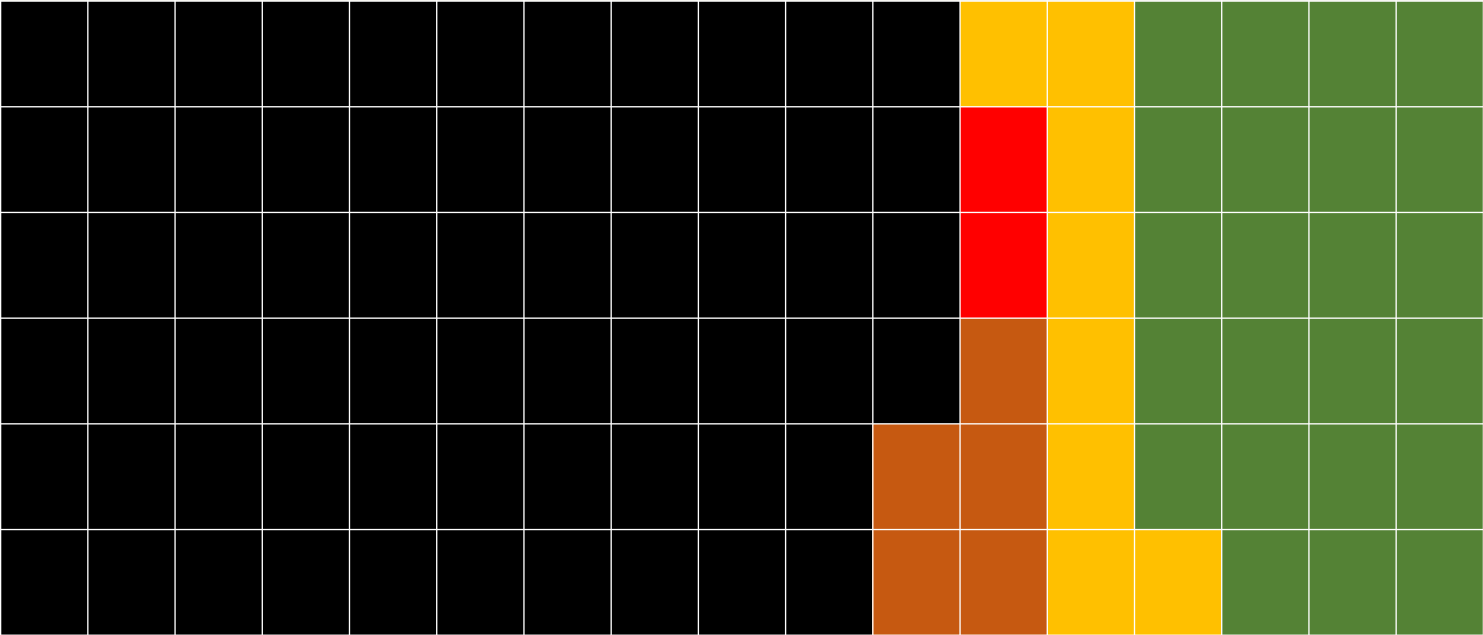
No access

Proprietary

Major differences

Minor differences

Comparable

PANELISTS

Name	Org name	Org type	Data policy	How is your organization contributing to the open evaluation data movement?
Neeta Goel	3ie	NGO	Yes	3ie mandates pre-registration and open access to data and pubs. Replicates results before releasing final payments.
Lorena Quintero	Gov of Colombia (GoC)	Gov	Yes	The GoC publishes data from every eval, following internal protocols of verification, replication, and anonymization.
Romina Cavatassi	IFAD	IO	No	IFAD has set up a task force to discuss data policy as the intention is to move towards data sharing and transparency.
Sudharsanam Balasubramaniam	CIFF	Funder	Yes	CIFF has an open data policy. All eval data are freely accessible for re-analysis.

AUDIENCE POLLING

- bit.ly/2J7o9kf
- Think of a question

AUDIENCE POLLING

PollEv.com/surveys/nEDYnG01Q/web

QUESTION #3

Do you see a role for replication files/open data in policymaking?

PANELISTS

Name	Org name	Org type	Data policy	How is your organization contributing to the open evaluation data movement?
Neeta Goel	3ie	NGO	Yes	3ie mandates pre-registration and open access to data and pubs. Replicates results before releasing final payments.
Lorena Quintero	Gov of Colombia (GoC)	Gov	Yes	The GoC publishes data from every eval, following internal protocols of verification, replication, and anonymization.
Romina Cavatassi	IFAD	IO	No	IFAD has set up a task force to discuss data policy as the intention is to move towards data sharing and transparency.
Sudharsanam Balasubramaniam	CIFF	Funder	Yes	CIFF has an open data policy. All eval data are freely accessible for re-analysis.

QUESTION #4

What are one to two recommendations would give to funders of evaluations around maximizing the value of their investments?

PANELISTS

Name	Org name	Org type	Data policy	How is your organization contributing to the open evaluation data movement?
Neeta Goel	3ie	NGO	Yes	3ie mandates pre-registration and open access to data and pubs. Replicates results before releasing final payments.
Lorena Quintero	Gov of Colombia (GoC)	Gov	Yes	The GoC publishes data from every eval, following internal protocols of verification, replication, and anonymization.
Romina Cavatassi	IFAD	IO	No	IFAD has set up a task force to discuss data policy as the intention is to move towards data sharing and transparency.
Sudharsanam Balasubramaniam	CIFF	Funder	Yes	CIFF has an open data policy. All eval data are freely accessible for re-analysis.

INCENTIVIZING REPLICATION FILES

Open data challenges

3ie – Government of Colombia
- 3ie and GoC data sets
- 2 prizes (class gift certificates)
- Internal and external review committee
- Winners announced on website, presented for GoC

Millennium Challenge Corporation
- MCC data sets
- Student funding for graduate school
- Internal (?) review committee
- Winner presented for MCC, job market paper

ADDITIONAL RESOURCES

Replication files
- [Data collection and file creation guide](#)
- [Best practices for creating reusable Dryad data packages](#)
- [Replication and the Manufacture of Scientific Inferences: A Formal Approach](#)"

Archiving data
- [Metadata standards and models](#)
- [Statistical disclosure control for microdata: A practice guide](#)

Open data studies
- [Open data impact when demand and supply meet](#)
- [Promoting an open research culture](#)
- [Push button replication: Is impact evaluation evidence for international development verifiable?](#)