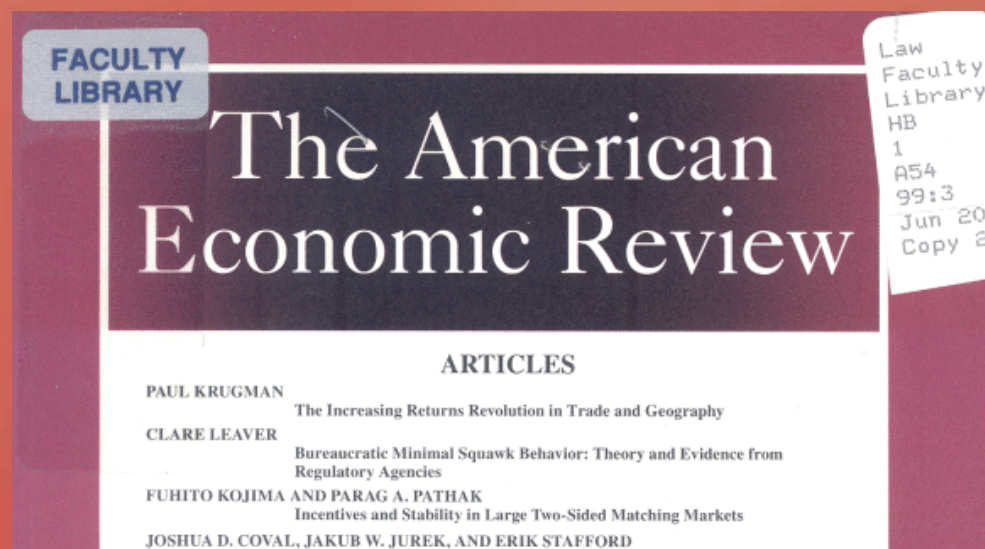


EDaWaX

European Data Watch Extended



Research Data Management in Economics Journals – Data Policies and Data Description as prerequisites of reproducible research

Sven Vlaeminck / Ralf Toepfer | Leibniz-Information Centre for Economics (ZBW)
2013-5-31 | IASSIST 2013 | Cologne | Germany

Funded by



EDaWaX

European Data Watch Extended

RatSWD.

Rat für Sozial- und
WirtschaftsDaten

ZBW

Leibniz-Informationszentrum
Wirtschaft
Leibniz Information Centre
for Economics

Table of Contents

- > Project background, aims and expected output
- > Some results of the project's analyses phase.
 - > Data sharing among economists
 - > Replicable research in economics journals
 - > Journals' infrastructures for managing research data
- > Our approach to facilitate metadata creation for researchers.

Project Background

Facilitate Replication of published empirical Findings in Economics.

"If the empirical basis for an article or book cannot be reproduced, of what use to the discipline are its conclusions? What purpose does an article like this serve?"

Gary King, 1995

Do you trust in applied economic research?

- > Dewald et al. (1986) tried to replicate 54 articles of the *Journal of Money, Credit, and Banking (JMCB)*.
 - They succeeded two times (3.7%).
 - > McCullough et al. (2006) tried to replicate 62 papers of the *JMCB*.
 - They were able to replicate 14 (22.6%).
 - > McCullough et al. (2008) tried to replicate 117 articles of the *Federal Reserve Bank of St. Louis Review*.
 - They were able to replicate 9 (7.7%).
- Often, published economic research is not replicable.

Background of the Project (II):

- > Explanations:
 - > Lack of incentives for sharing data / lack of recognition and credit
 - > Infrequent implementation of data availability policies by economics journals
 - > Infrastructure for publication-related research data management is rarely available
- > EDaWaX addresses these three challenges

The EDaWaX-Project

- > EDaWaX analysed...
 - ...journals' data policies and made recommendations for those guidelines.
 - ...analysed the data sharing behaviour of economists.
 - ... data centres and made recommendations where to host a publication related data archive.
- > EDaWaX...
 - ...used existing metadata schema to derive three „levels“ of metadata to describe data and publications.
- > Based on these findings we intend to develop & to implement a publication-related data archive for an economic journal (using CKAN).

Data Sharing among Economists

The current state of data sharing in the profession.

Funded by
DFG

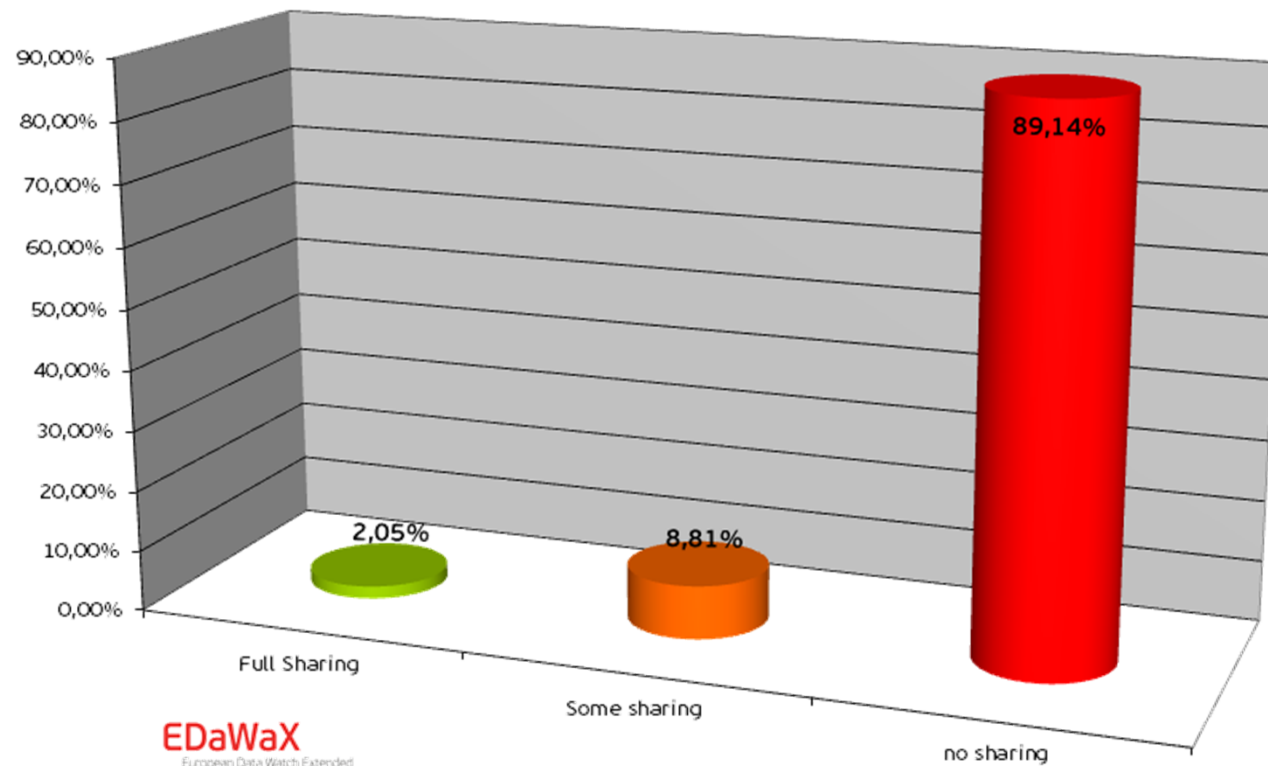
EDaWaX
European Data Watch Extended

RatSWD.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

The Status Quo among Economists...

Data Sharing among Economists
(n=488)



Source: Andreoli-Versbach/Mueller-Langer (2013): Open Access to Data: An Ideal Professed but not Practised

...is not to share their data

Replicability and research data management in economics journals

"Most results published in economics journals cannot be subjected to verification, even in principle, because authors typically are not required to make their data and code available for verification."¹

McCullough, McGeary, Harrison (2006)

Funded by

DFG

EDaWaX

European Data Watch Extended

RatSWD.

ZBW

Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Replicability in economics Journals

141

Journals
in our
Sample

We built a sample of 141 economics journals to evaluate the amount of journals equipped with data policies.

EDaWaX

European Data Watch Extended

RatSWD.

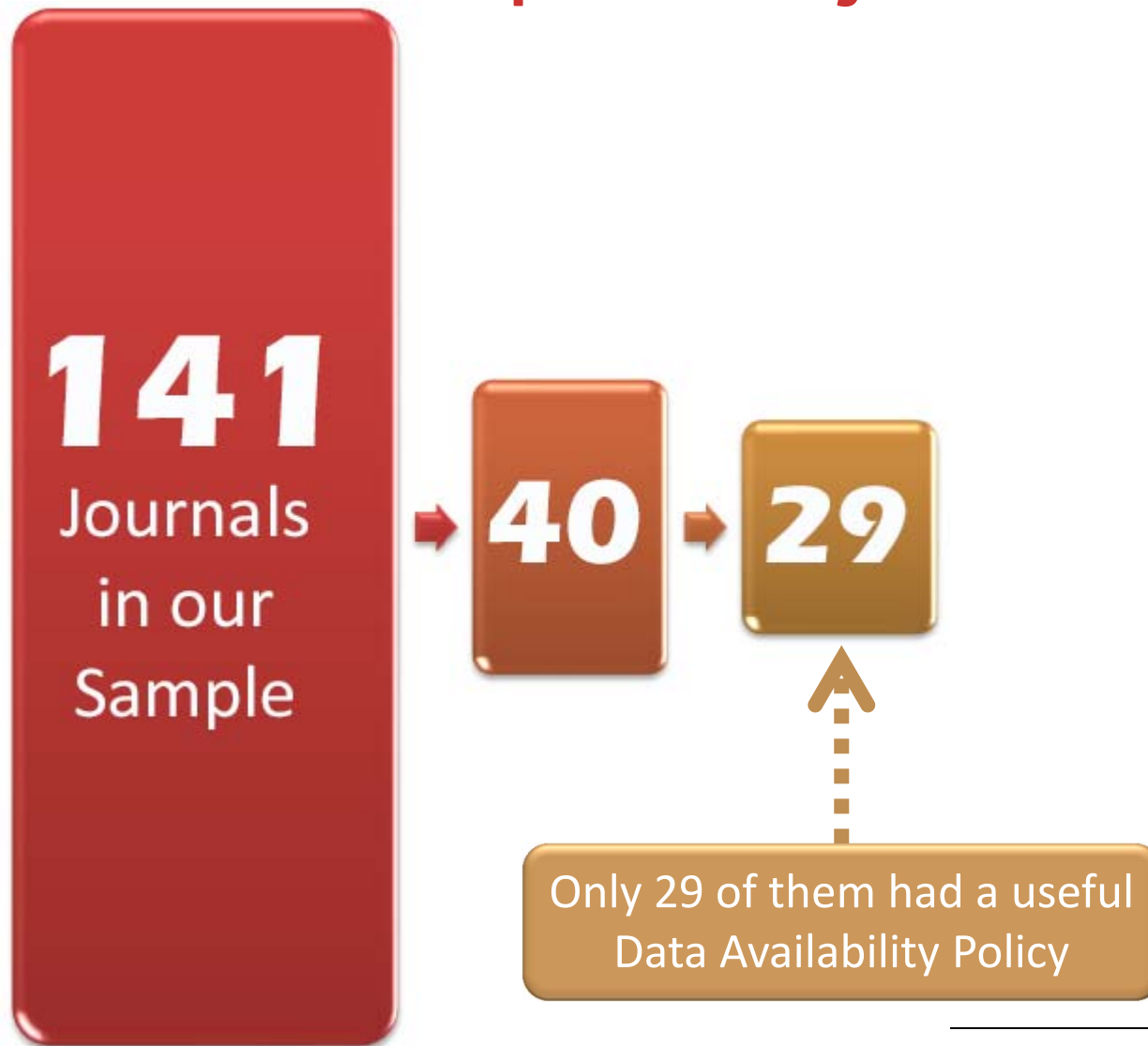
ZBW

Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Replicability in economics Journals



Replicability in economics Journals



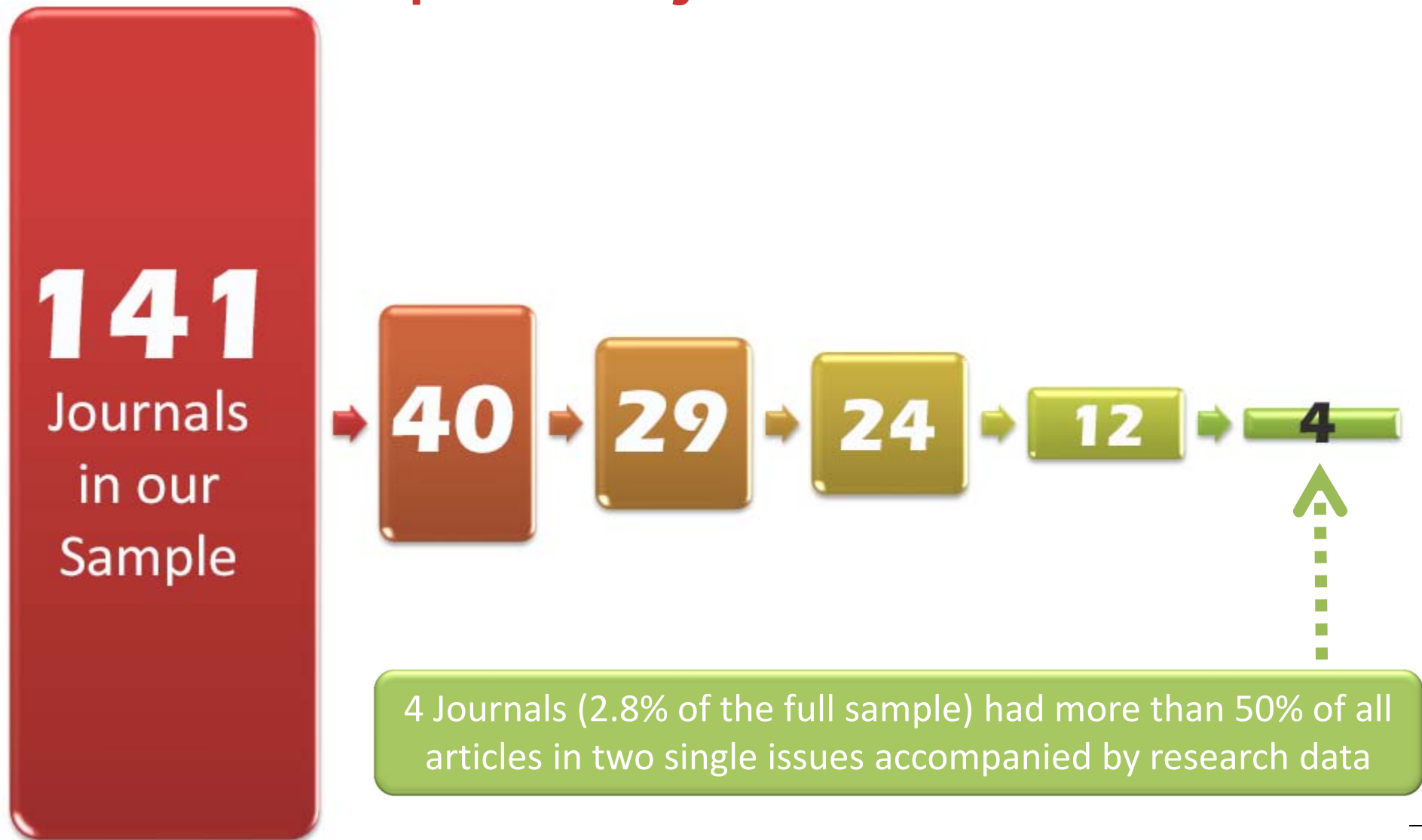
Replicability in economics Journals



Replicability in economics Journals



Replicability in economics Journals



The journals' infrastructure to provide research data

Funded by
DFG

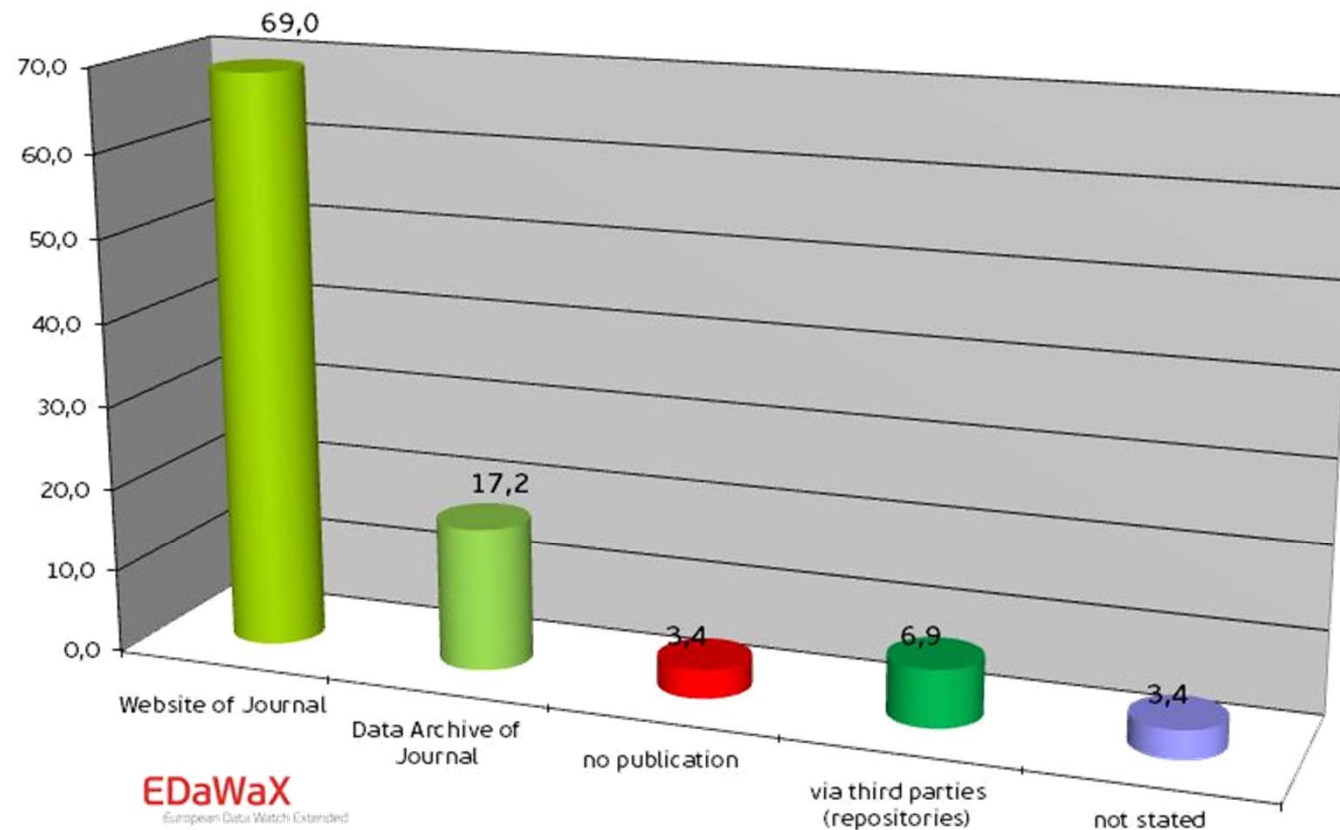
EDaWaX
European Data Watch Extended

RatSWD.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

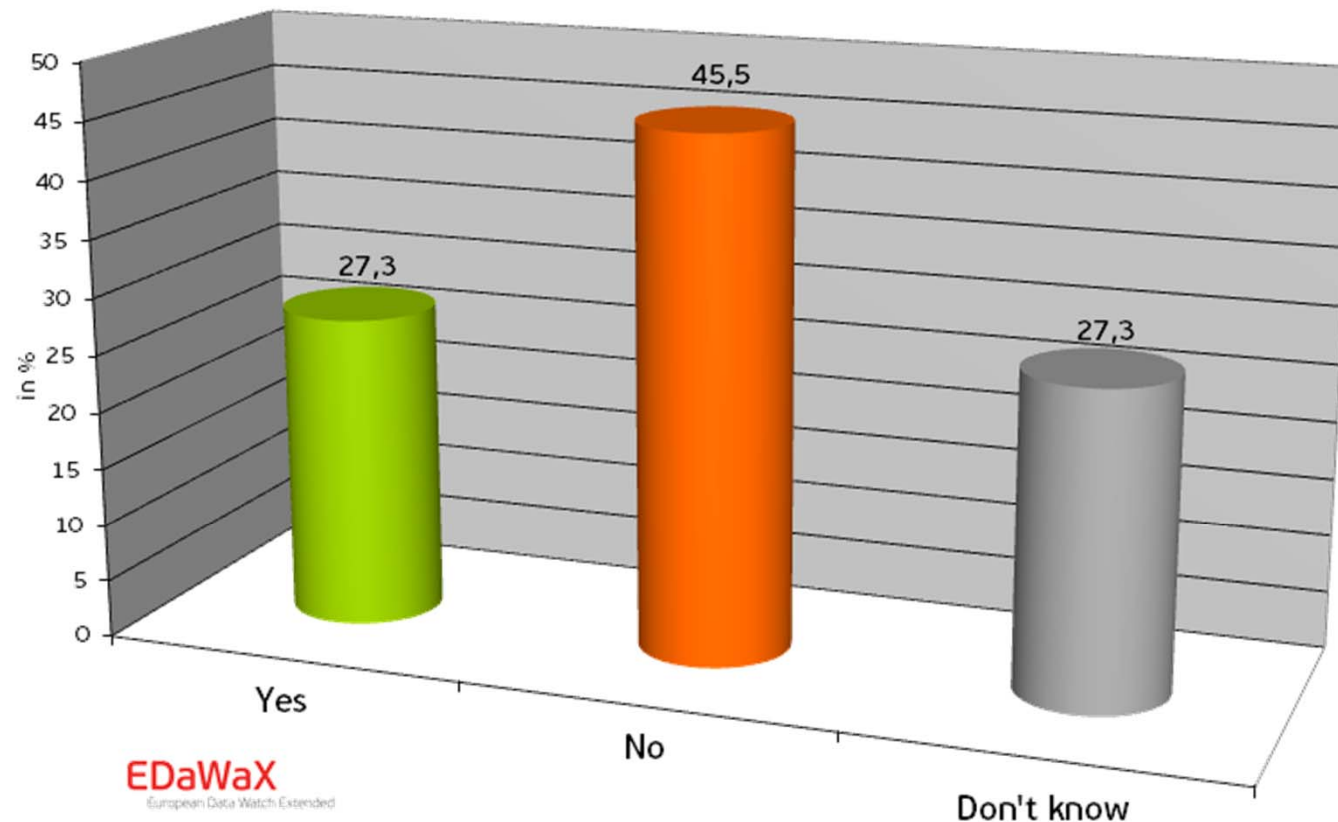
The current e-infrastructure is not sufficient...

How do Journals publish Research Data and Supplementary Materials?
(n=29)



Often, additional metadata is not created...

Creation of additional metadata and/or persistent identifiers
for publication-related research data
(n=11)



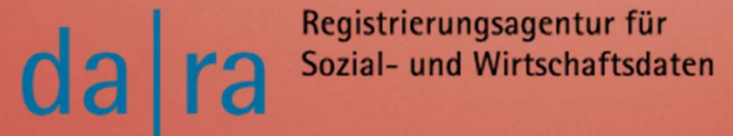
Summary: Publication-related research data management in economics journals

- > Economists rarely share their research data with others.
- > Only a few journals own usefull data availability policies AND enforce data availability.
- > Research data often is available in form of zip-files only and it is accessible mainly via the publisher's website.
- > Mostly no specific metadata / no persistent identifiers added
- > Data sharing and research data management in economics journals are still at its early stages!

How can we increase data availability and reproducible research?

- > Build a sound, usefull infrastructure to support RDM
- > Improve data policies
- > Rise awareness that data sharing and working up data is an important (scientific) task.
- > Support incentive-mechanisms, e.g.
 - make data citable (-> DataCite)
 - include research data in our disciplinary portals.
 - implement metrics and usage statistics.
- > But: For doing so, we need additional metadata! Who can create it?

Our approach to facilitate metadata creation for researchers.



Funded by
DFG

EDaWaX
European Data Watch Extended

RatSWD.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

How much documentation do we need?

And who is going to create it?

- > For replication purposes it is necessary to use adequate metadata schema (e.g. DDI3)
- > ...but extensive schema are not accepted by researchers. What to do?

Our approach:

- > We defined some „levels“ and functionalities/purposes of metadata.
 - > Researcher chooses metadata „level“ and functionalities.
 - > Important to point out advantages and limitations of each „level“ of data description!
-

Metadata for publication-related research data

> We identified four „levels“ of metadata – three of them are available in our pilot application:

1. Level: *Ensure citability (DataCite/da/ra)* -
 - 9 metadata fields
2. Level: *Support findability (da/ra):*
 - 26 metadata fields
3. Level: *Ensure linkability (da/ra) :*
 - up to 100 fields
4. *Level: Reproducibility (DDI3):*
 - up to 846 fields

Vision

- > *“Making it convenient for scientists to describe, deposit, and share their data and to access data from others, plus promulgating the best data practices through education and awareness will help the future of science as well as the future of data preservation.”*

(Tenopir 2011:20)

Thank you very much for your attention!

Contact:

Sven Vlaeminck

s.vlaeminck@zbw.eu

Ralf Toepfer

r.toepfer@zbw.eu

<http://www.edawax.de> (in English)

Funded by

DFG

EDaWaX

European Data Watch Extended

RatSWD.

ZBW

Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics