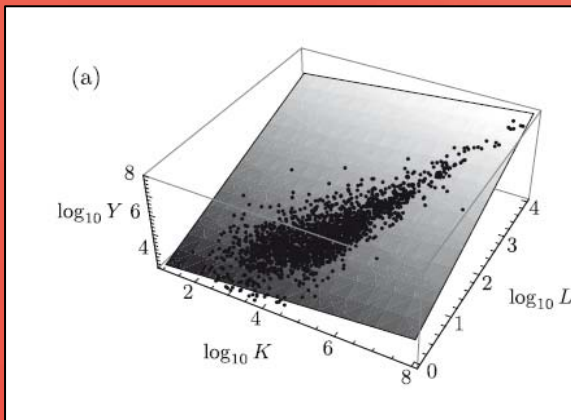


EDaWaX

European Data Watch Extended



Source: *Economics-EJournal*

[The Welfare Effects of Bundling in Multichannel Television Markets](#) (pp. 643-85)

Gregory S. Crawford and Ali Yurukoglu

[Abstract/Tools](#) | [Full-Text Article](#) | [Download Data Set](#) | [Online Appendix](#)

[The Use of Full-Line Forcing Contracts in the Video Rental Industry](#) (pp. 686-719)

Katherine Ho, Justin Ho and Julie Holland Mortimer

[Abstract/Tools](#) | [Full-Text Article](#) | [Download Data Set](#) | [Online Appendix](#)

[Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World](#) (pp. 720-49)

Drew Fudenberg, David G. Rand and Anna Dreber

[Abstract/Tools](#) | [Full-Text Article](#) | [Download Data Set](#) | [Online Appendix](#)

Source: *American Economic Review*



Source: Dataverse-Network

Data Management in Scholarly Journals and possible Roles for Libraries – Some Insights from EDaWaX

Sven Vlaeminck | Leibniz-Information Centre for Economics (ZBW)

2012-6-27 | LIBER 41st Annual Conference | Tartu | Estonia

Table of Contents

- > Project background, aims and expected output
- > Data availability policies in economic journals
 - > Status Quo – do you trust in economic research?
 - > Enabling replications – requirements for data policies
 - > Some results of our study
- > Linking data and publications – roles and responsibilities

Project Background

Enabling Replications of empirical Findings in Economics.

The EDaWaX Project | www.edawax.de

The Background of our Project:

- > Increasing number of empirical/econometric publications in economics
- > No means to replicate the results of economic research
 - > Incentives for researchers to share „their“ data are missing
 - > Infrequent implementation of data availability policies by scholarly journals
 - > E-infrastructure for publication-related research data is not available
- > Replication is a cornerstone of the scientific method!

Project Phases and Outcome

Project Phase 1: ANALYSIS

- Evaluation of Incentives for Economists to publish Research Data.
- Analysis of Data Policies of 140 Economic Journals.
- Evaluation of Hosting Options in Research Data Centers.

Project Phase 3: IMPLEMENTATION & EVALUATION

- Implementation of the Pilot-Application.
- Evaluation of Project Results by Editors of Economic Journals and Research Data Centers.
- Adjustments and Evaluation by Authors of „Schmollers Jahrbuch“

Project Phase 2: CONCEPTION

- Development of a Metadata Schema for Description and Labeling of Research Data
- Generation of functional Requirements for a Pilot-Application based on the Results of Project Phase 1

Project output:

- > Implementation of a data archive for a scholarly journal.
- > Incentive schema for authors to publish research data.
- > Recommendations for data policies.
- > Recommendation for hosting options for research data.
- > Metadata schema for the description and labelling of research data.

Data Availability Policies in Economic Journals

"Results published in economic journals are accepted at face value and rarely subjected to the independent verification that is the cornerstone of the scientific method. 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."

B.D. McCullough

Do you trust in Economic Research?

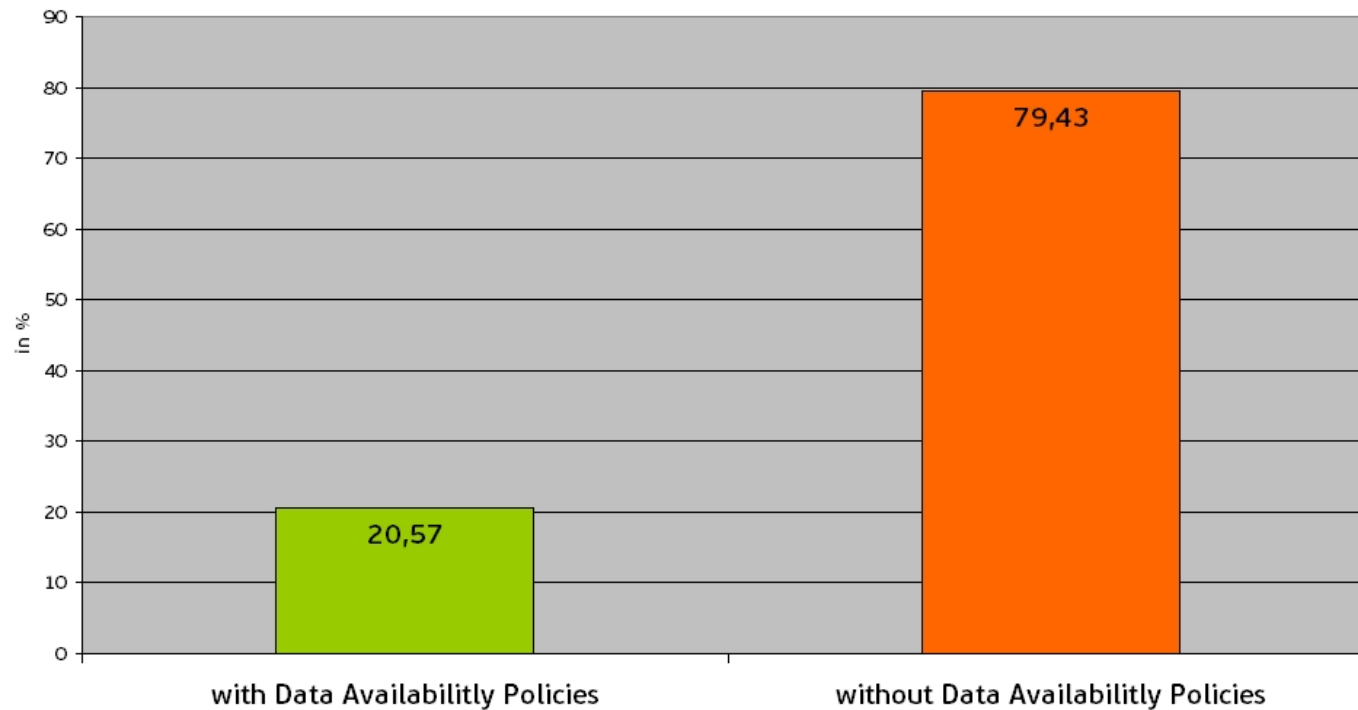
- > Dewald et al. (1986) tried to replicate 54 articles of the *Journal of Money, Credit, and Banking (JMCB)*.
 - They succeeded two times (3.7%).
 - > McCullough et al. (2006) tried to replicate 62 papers of the *JMCB*.
 - They were able to replicate 14 (22.6%).
 - > McCullough et al. (2008) tried to replicate 117 articles of the *Federal Reserve Bank of St. Louis Review*.
 - They have been successful with 9 (7.7%).
- ...often, economic research is not replicable.
But: What do economists need for successful replications?

(Some) Requirements for Data Policies

- > Data Policies that aim to ensure the replicability of publication-related research data in economics have to...
 - > ...be mandatory.
 - > ...pledge authors to provide datasets, code and programs as well as descriptions of the data and variables (-> data dictionary).
 - > ...assure that the data is provided prior to the publication of an article.
 - > ...make the data and other materials available to interested researchers.
 - > ...have defined rules if proprietary or confidential data was used.
- ➔ These requirements have been used to analyze the data policies of 141 economic journals.

Journals with Data Availability Policies

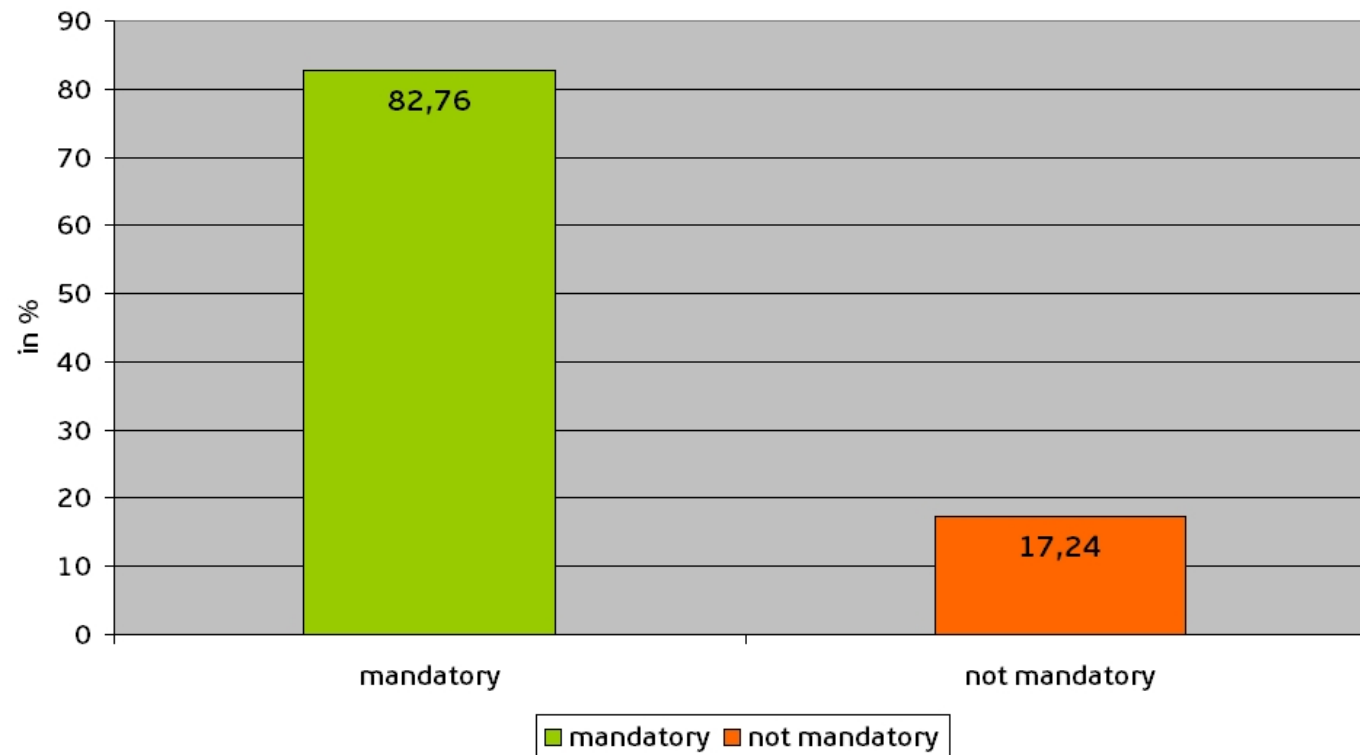
Journals with Data Availability Policies in our Sample
(n=141)



> More than 20% (29) of all journals in our sample own a data availability policy

The Number of mandatory Data Policies

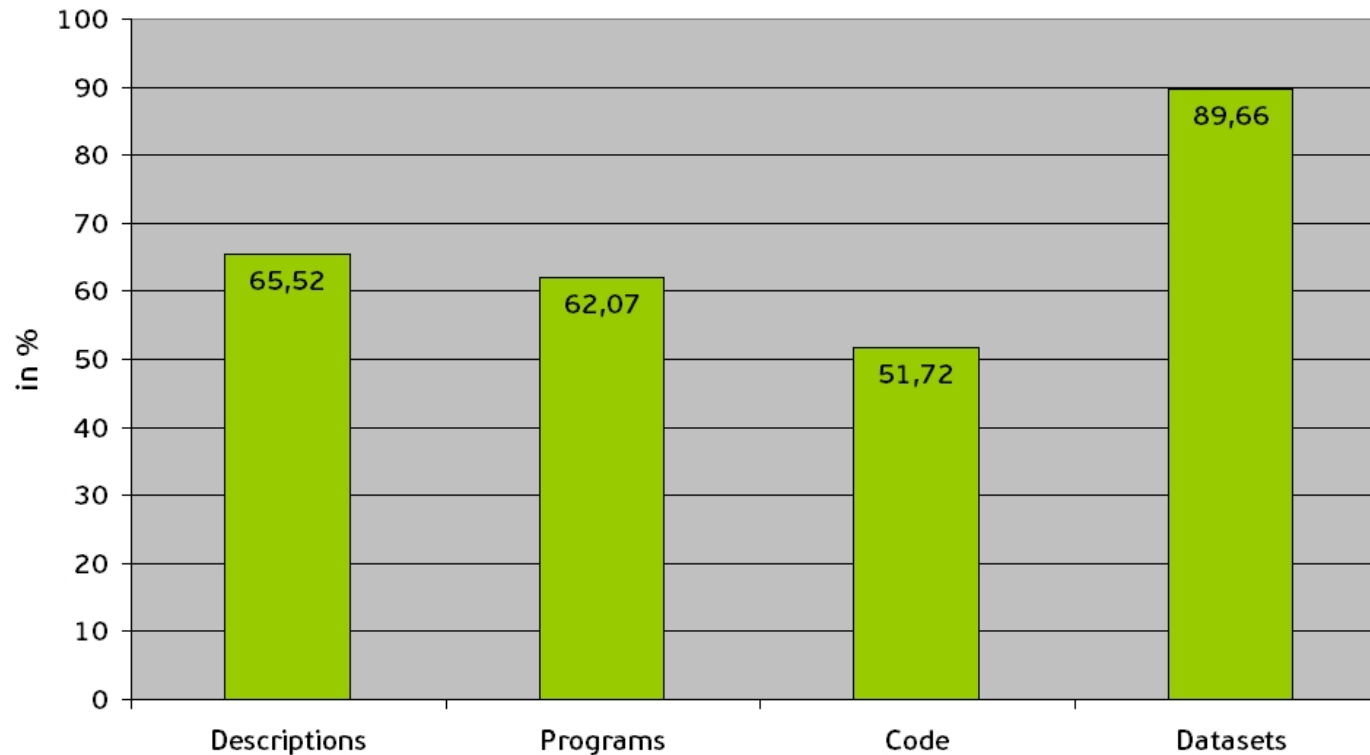
Mandatory or not? Data Availability Policies in our Sample (n=29)



> The big majority of the 29 data policies is mandatory

What kinds of data authors have to provide...

Data Availability Policies: What Kinds of Data have to be provided ? (n=29)

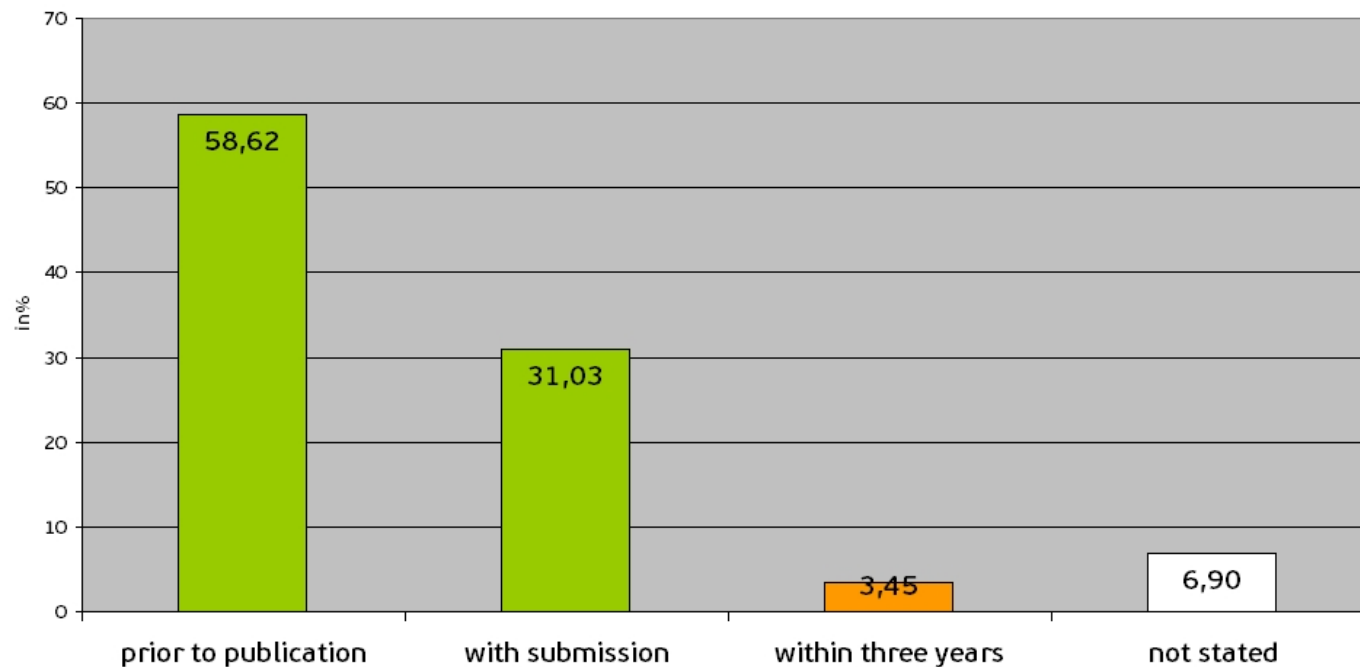


- > Nearly all policies pledge authors to provide datasets. Descriptions and programs are required by $\sim \frac{2}{3}$. Code by $\sim 50\%$ only.

When do authors have to provide their data?

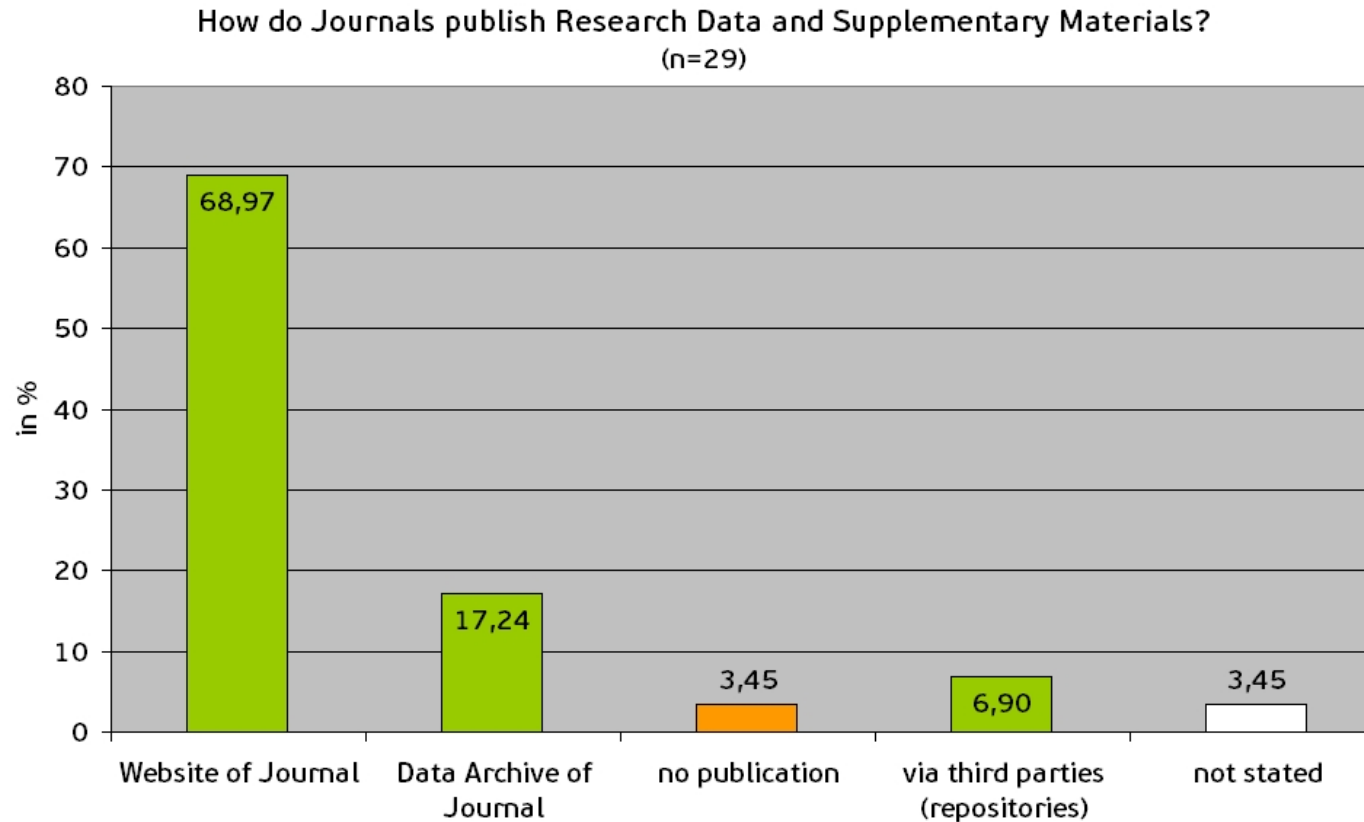
Journals with Data Availability Policies:
Point of Time to provide Data and other Materials

(n=29)



> ~ 90% of the policies pledge authors to provide the data prior to publication

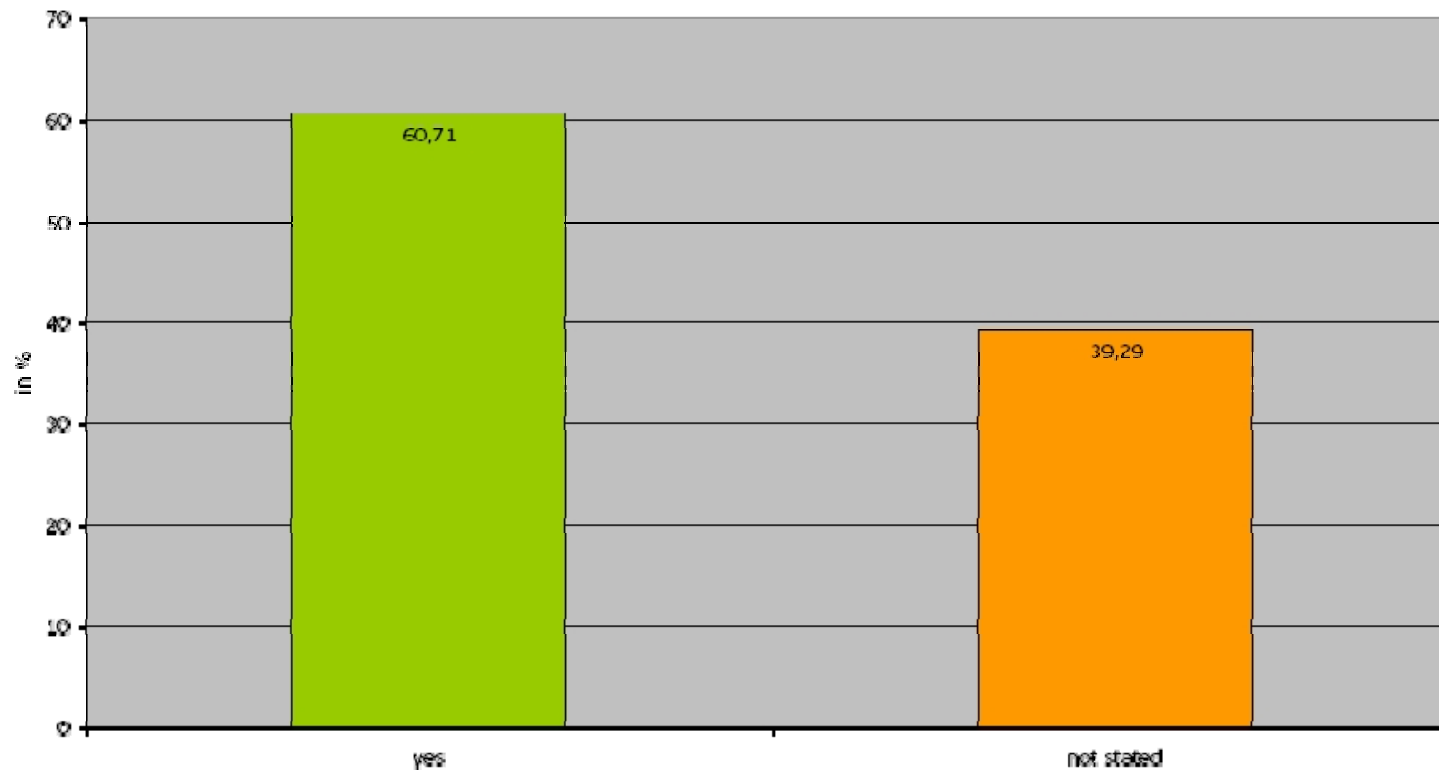
How do Journals publish the Data?



- > Most often data is published in the 'supplementary material' section of the journal's website

Defined Procedure in Case of proprietary Data

Do Journals own a defined Procedure in Case of proprietary or confidential Data?
(n=29)



> ~ 60% of the policies implemented procedures for research based on proprietary data

Results of our Study

- > Percentage of journals with data availability policy still is small – but it is rising.
- > Policies differ massively in quality and extent.
- > Majority of the policies analyzed are mandatory.
- > Almost all journals require the submission of datasets, prior to publication of an article.
- > Less than $\frac{2}{3}$ of all policies require the submission of programs and descriptions of the data.
- > Only 50% of the data policies require the submission of code.
- > $>\frac{3}{4}$ allowing exceptions to the data policy...
- > ...but only 60% have rules for authors that used proprietary data.

Gap - Analysis

- > We need...
 - Incentives for researcher to provide „their“ data.
 - A ‚cultural change‘ in Economics: Replications and the provision of data are crucial for the progress of science.
 - Standards - what kind of data has to be provided in which granularity for enabling replications?
 - An improved e-infrastructure for
 - ...providing data and materials to interested researchers.
 - ...making data citable, searchable and reusable.
 - ...safeguarding and long-term preservation of these data

Linking Data & Publications

...a future task for Research Libraries ?!

Roles, Rights, Responsibilities

- > Major stakeholders involved :
 - Researcher (as creator/user) of data
 - Editors of scholarly journals
 - Publishers
 - Data centres
 - Research libraries
- > For the goal of linking data and publications, all of the stakeholders have their roles and responsibilities:



Researchers: Roles and Responsibilities



Role: creator of data

- > Meet standards for good practice.
- > Work up data for use by others.
- > Comply with journal's data policies.
- > Submit data to journal's data archive.
- > Submit core metadata.

Illustration: TheQuiltedLion / flickr.com

Editors: Roles and Responsibilities



Role: enforce data availability

- > Implement data policy for journal.
- > Monitor and enforce data availability.
- > Ensure data is stored in a trustworthy place or repository.
- > Negotiate with publisher to enable linking to journal's data archive

Illustration: MacKenzie Smith | MIT-Library

Publisher: Roles and Responsibilities



Role: link dataset to article

- > Link to research data to support publication standards.
- > Support uniform data citation standards.

Illustration : nowwiskie / flickr.com

Data Centres: Roles and Responsibilities



Role: curation of and access to data

- > Develop easy to use user front-ends to facilitate data submission.
- > Create technical metadata.
- > Manage data (and software) for the long-term.
- > Provide training for deposit.
- > Manage data access according to IPR.
- > Provide tools for re-use of data.
- > Creation of persistent identifiers.

Illustration: J__st / flickr.com

Research Libraries: Roles and Responsibilities



*Role: cataloguing, retrieval,
content acquisition*

- > Develop easy to use user front-ends to facilitate data submission.
- > Create of persistent identifiers.
- > Provide training for deposit.
- > Create further descriptive and administrative metadata.
- > Provide interoperable metadata (schemata).
- > Content acquisition of datasets.
- > Cataloguing research data and publication.
- > Integrate research data in retrieval services and link data and publications.

Illustration: MacKenzie Smith | MIT-Library

Thank you very much for your attention!

...any questions / comments?!

Contact:

s.vlaeminck@zbw.eu | www.edawax.de