

## Leibniz - Bibliotheksverbund Forschungsinformation



„Wirtschaftswissenschaftliche Forschungsergebnisse replizierbar machen – das Projekt EDaWaX“

101. Deutscher Bibliothekartag | 24.5.2012 | CCH Hamburg

Sven Vlaeminck | Leibniz-Informationszentrum Wirtschaft (ZBW)

**EDaWaX**

European Data Watch Extended



## Inhaltsübersicht:

- > Projekthintergrund
- > Vorgehen und Zwischenstand Projektphase 1
  - > Vorgehen und Zwischenergebnisse „ökonomische Anreizanalyse“
  - > Vorgehen und Zwischenergebnisse „Hostingoptionen bei FDZs“
- > Data Policies in wirtschaftswissenschaftlichen Fachzeitschriften
  - > Status Quo – trauen Sie wirtschaftswissenschaftlicher Forschung?
  - > Anforderungen an Data Policies zur Gewährleistung von Replikationen
  - > Ergebnisse unserer Studie
- > Fazit

# Projekthintergrund

Wirtschaftswissenschaftliche Forschungsergebnisse replizierbar machen. Das Projekt EDaWaX | [www.edawax.de](http://www.edawax.de)

Gefördert durch:

**DFG**

**EDaWaX**  
European Data Watch Extended

**RatSWD.**  
Rat für Sozial- und  
Wirtschaftsdaten

**ZBW** Leibniz-Informationszentrum  
Wirtschaft  
Leibniz Information Centre  
for Economics

# Hintergründe des Projekts

- > Zunehmend mehr empirische/ökonometrische Studien in den Wirtschaftswissenschaften
- > Kaum Möglichkeiten zur Replikation von Ergebnissen
  - > fehlende Anreize für Forscher „ihre“ Daten zu publizieren
  - > Forschungsdaten-Policies in Fachzeitschriften sind noch selten
  - > Fehlende / im Aufbau befindliche Infrastruktur für publikationsbezogene Forschungsdaten
- > Replikationen sind ein zentrales wissenschaftliches Prinzip!

# Projektphasen & Projektziel



## IMPLEMENTIERUNG & EVALUIERUNG

- Entwicklung / Anpassung einer Pilotanwendung
- Evaluierung d. Projektergebnisse durch Fachcommunity / Hrsg.
- Anpassung d. Software an die Bedürfnisse der Community

## Konzeption

- (Weiter-) Entwicklung eines Metadatenschemas für Forschungsdaten (-> DDI).
- Konzeption und Evaluierung von Softwarelösungen zum Übermitteln von Forschungsdaten

## ANALYSE

- Analyse der Data Policies von WiWi-Fachzeitschriften
- ökonomische Anreizanalyse
- Analyse von Hosting-Optionen der Forschungsdatenzentren

**EDaWaX**

European Data Watch Extended

RatSWD.

**ZBW**

Leibniz-Informationszentrum Wirtschaft  
Leibniz Information Centre for Economics



# Vorgehen & Zwischenstand der 1. Projektphase

Vorgehen und erste Zwischenergebnisse unserer Untersuchungen...

Gefördert durch:

**DFG**

**EDaWaX**  
European Data Watch Extended

RatSWD.

**ZBW** Leibniz-Informationszentrum Wirtschaft  
Leibniz Information Centre for Economics

# Vorgehen „ökonomische Anreizanalyse“

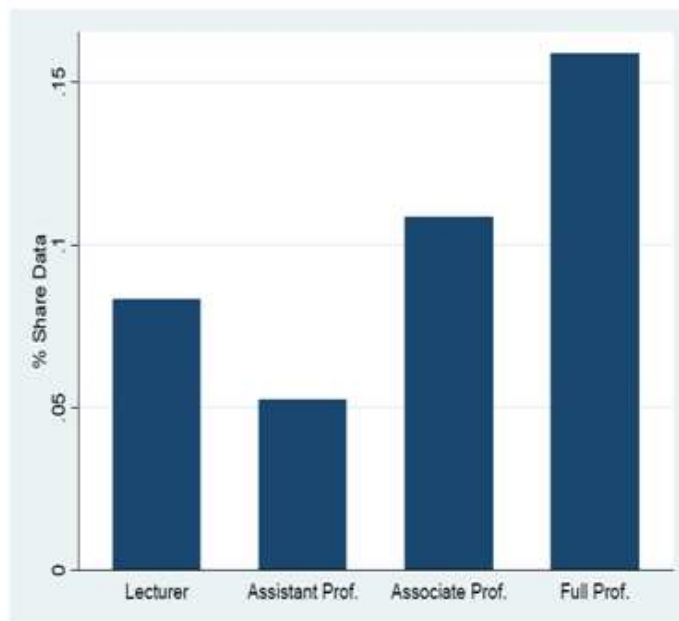


- > 3600 deutsche Wirtschaftswissenschaftler werden zu Anreizen befragt, ihre Daten zu teilen („Data Sharing“)
- > Auswertung der Data Sharing Praktiken von US- und europäischen Ökonomen (Desktop Research)
- > Online-Befragung der Herausgeber von 100 führenden Fachzeitschriften zu Erfahrungen mit Data Policies

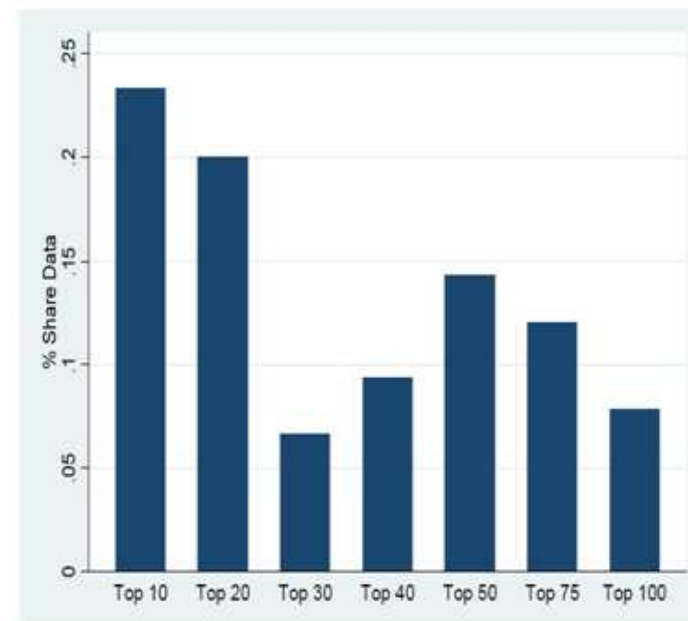
# Welche Wissenschaftler teilen ihre Daten?

Grafiken: Patrick Andreoli-Versbach (IMPRS-CI) / Frank Müller Langer (IMPRS-CI)

Preliminary results: Current Position



Preliminary results: Institutions' Rank



➤ Vor allem ordentliche Professoren und Top-Institute betreiben Data Sharing.



## Vorgehen „Analyse von Hostingoptionen “

- > Desktop Research bei 45 Datenzentren
- > Prüfung, ob Services für externe Daten angeboten werden
  - > Kosten ?
  - > Datenschutz & Sicherheit ?
  - > eingesetzte Infrastruktur und Metadaten ?
  - > Persistente Identifikatoren?
- > Vertiefende Analyse durch
  - > Interviews
  - > Onlinebefragung

# Data Policies von Fachzeitschriften in den Wirtschaftswissenschaften

"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

Gefördert durch:

**DFG**

**EDaWaX**

European Data Watch Extended

RatSWD.

**ZBW**

Leibniz-Informationzentrum Wirtschaft  
Leibniz Information Centre for Economics

# Trauen Sie wirtschaftswissenschaftlicher Forschung?

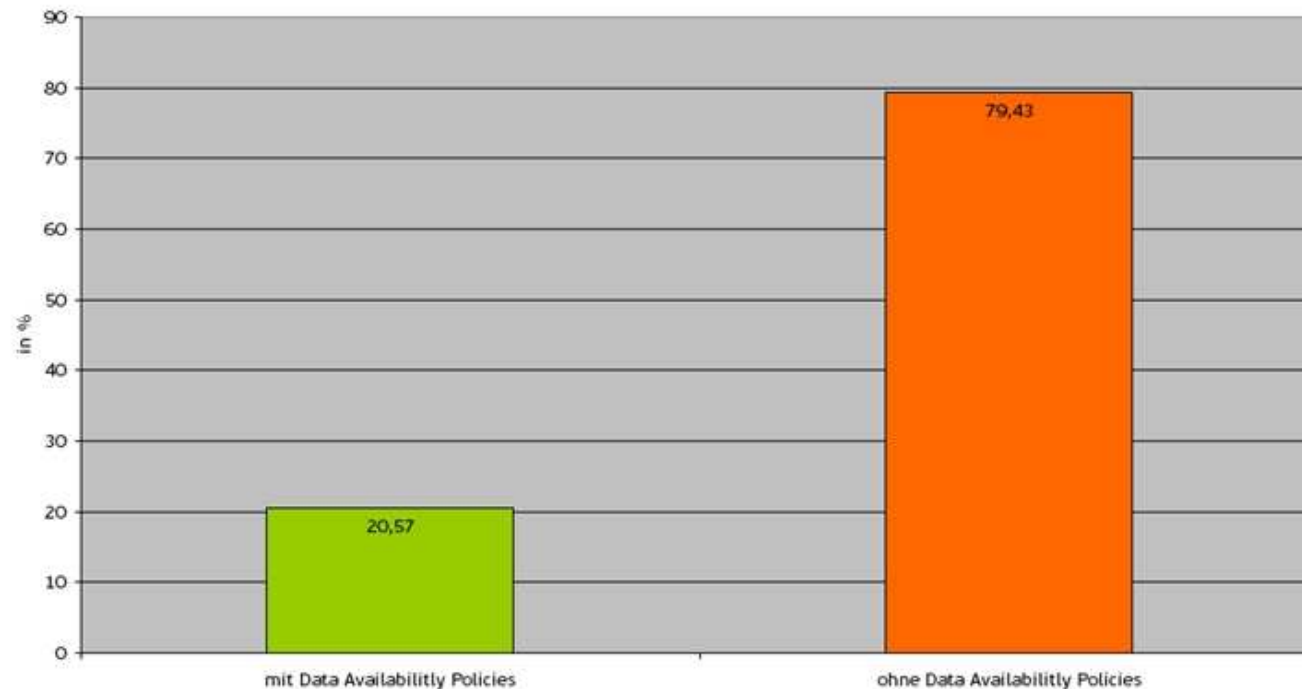
- > Dewald et al. (1986) versuchten 54 Artikel des *Journal of Money, Credit, and Banking (JMCB)* zu replizieren.
  - Es gelang bei zweien (3,7%).
- > McCullough et al. (2006) versuchten 62 Veröffentlichungen des *JMCB* zu replizieren.
  - Sie konnte 14 (22,6%) replizieren.
- > McCullough et al. (2008) versuchten 117 Artikel der *Federal Reserve Bank of St. Louis Review* zu replizieren.
  - Es gelang Ihnen bei 9 (7,7%).
- ...wirtschaftswissenschaftliche Forschung ist oft nicht replizierbar.
- ...doch was benötigen Ökonomen für erfolgreiche Replikationen?

# Anforderungen an Data Policies

- > Richtlinien zur Übermittlung publikationsbezogener Forschungsdaten zur Gewährleistung von Replikationen müssen...
  - > ... verbindlich sein.
  - > ... neben Datensätzen auch Berechnungscode und zur Simulation genutzten Programme einfordern.
  - > ... Beschreibungen der übermittelten Daten bieten (-> Data Dictionaire).
  - > ...die Daten vor der Veröffentlichung des Artikel einfordern.
  - > ...die von ihnen eingeforderten Daten bereitstellen.
  - > ...über ein definiertes Vorgehen im Fall von proprietären Daten verfügen.
- > Prüfung von 141 wirtschaftswissenschaftlichen Fachzeitschriften auf Verhandensein und Qualität von Data Availability Policies

# Zeitschriften mit Data Availability Policies

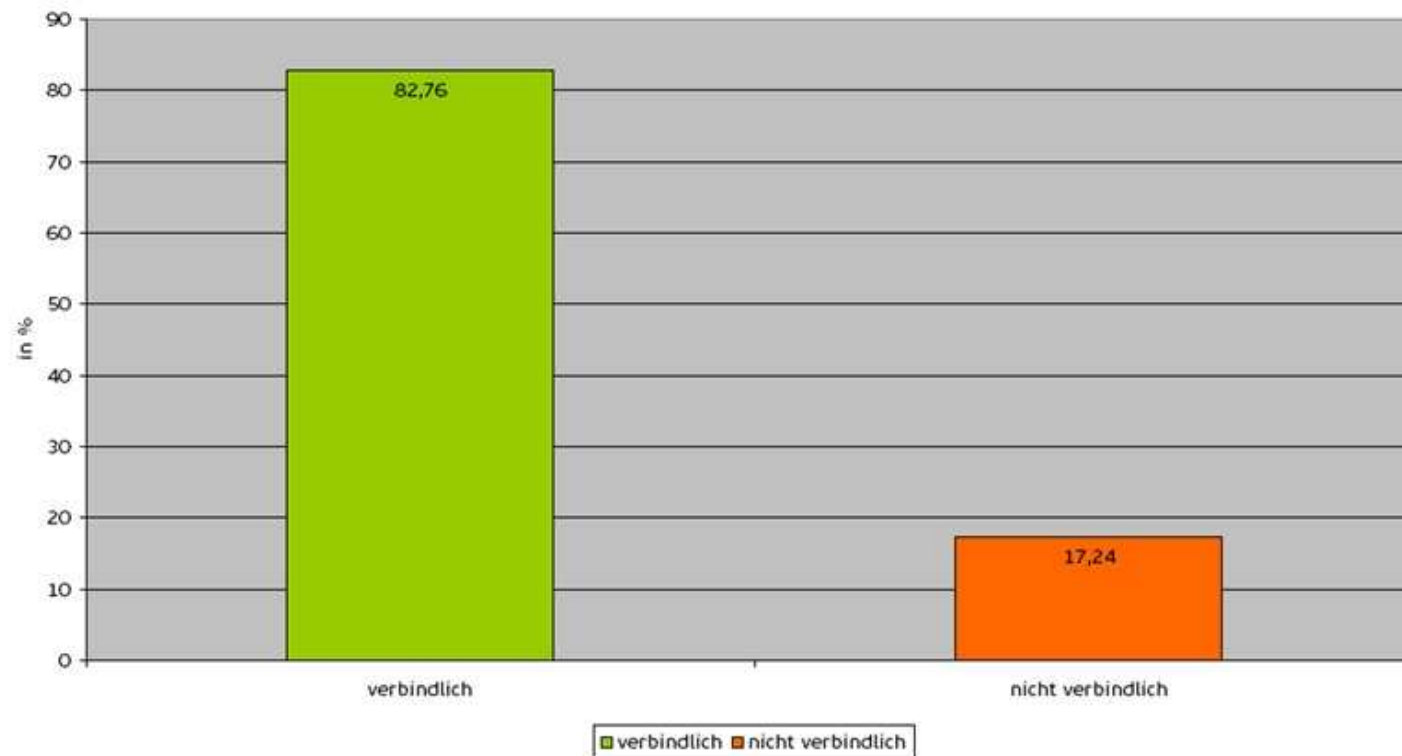
Zeitschriften mit Data Availability Policies im Untersuchungssample  
(n=141)



>>20% (29) der untersuchten Zeitschriften verfügen über eine Data Availability Policy

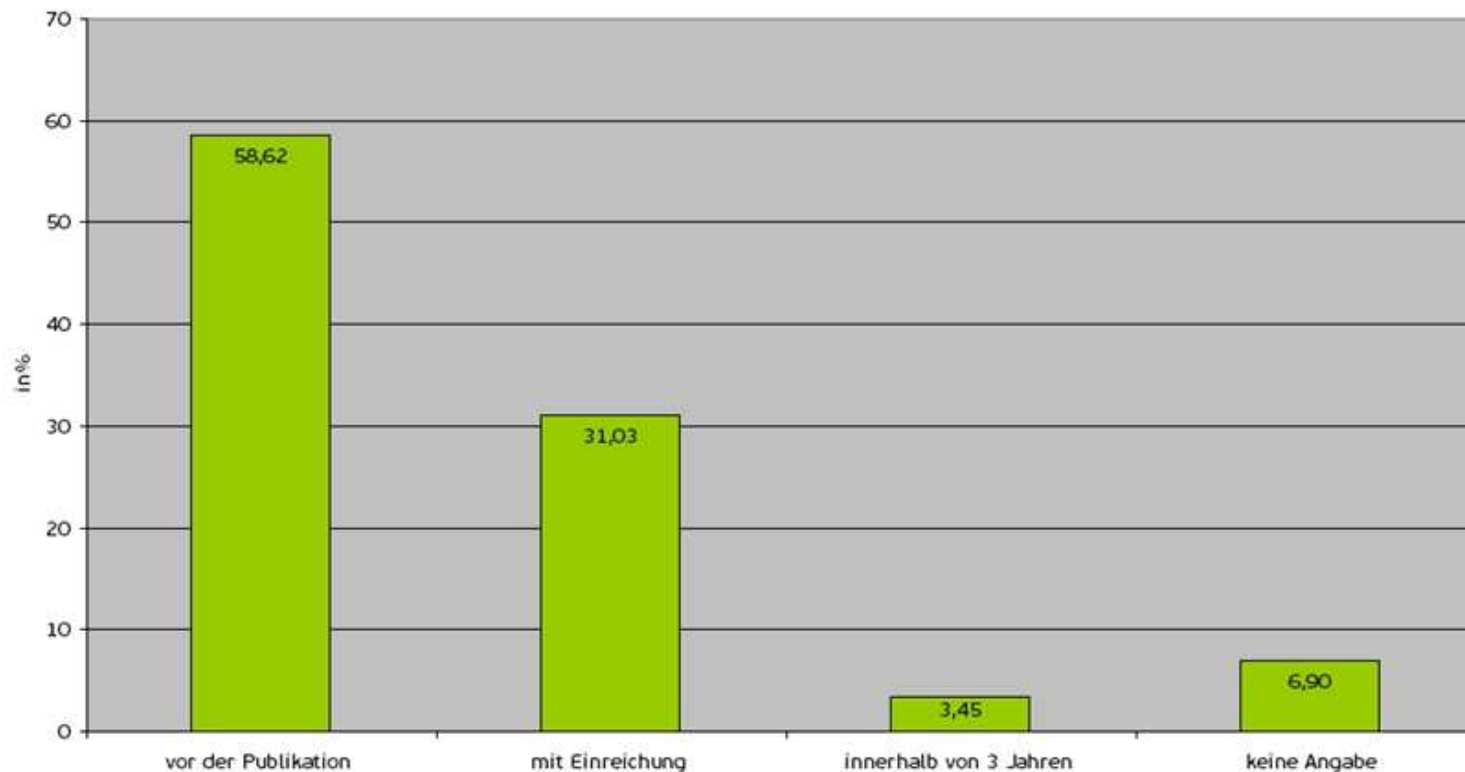


# Die Verbindlichkeit der Richtlinien für Autoren



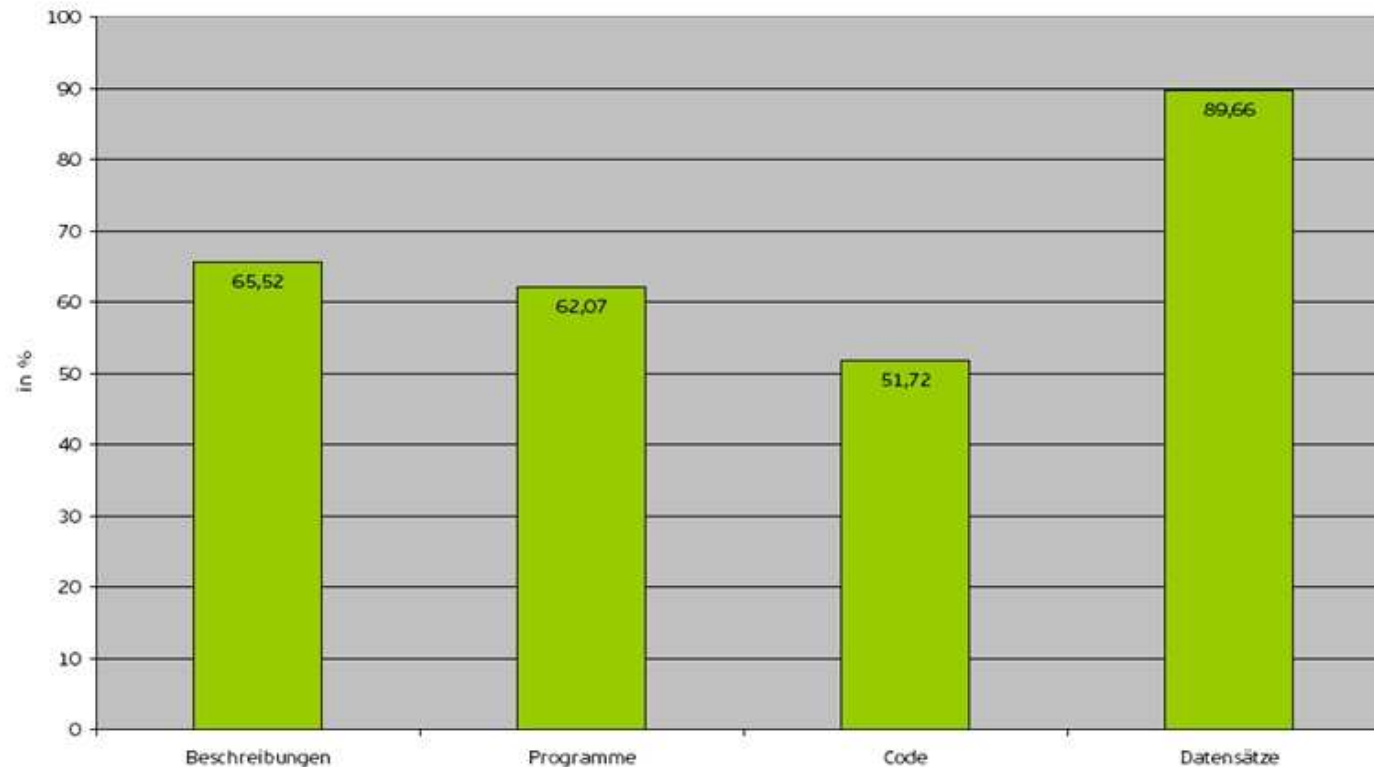
> Die übergroße Mehrheit der untersuchten 29 Richtlinien ist verpflichtend

# Zeitpunkt der Datenübermittlung



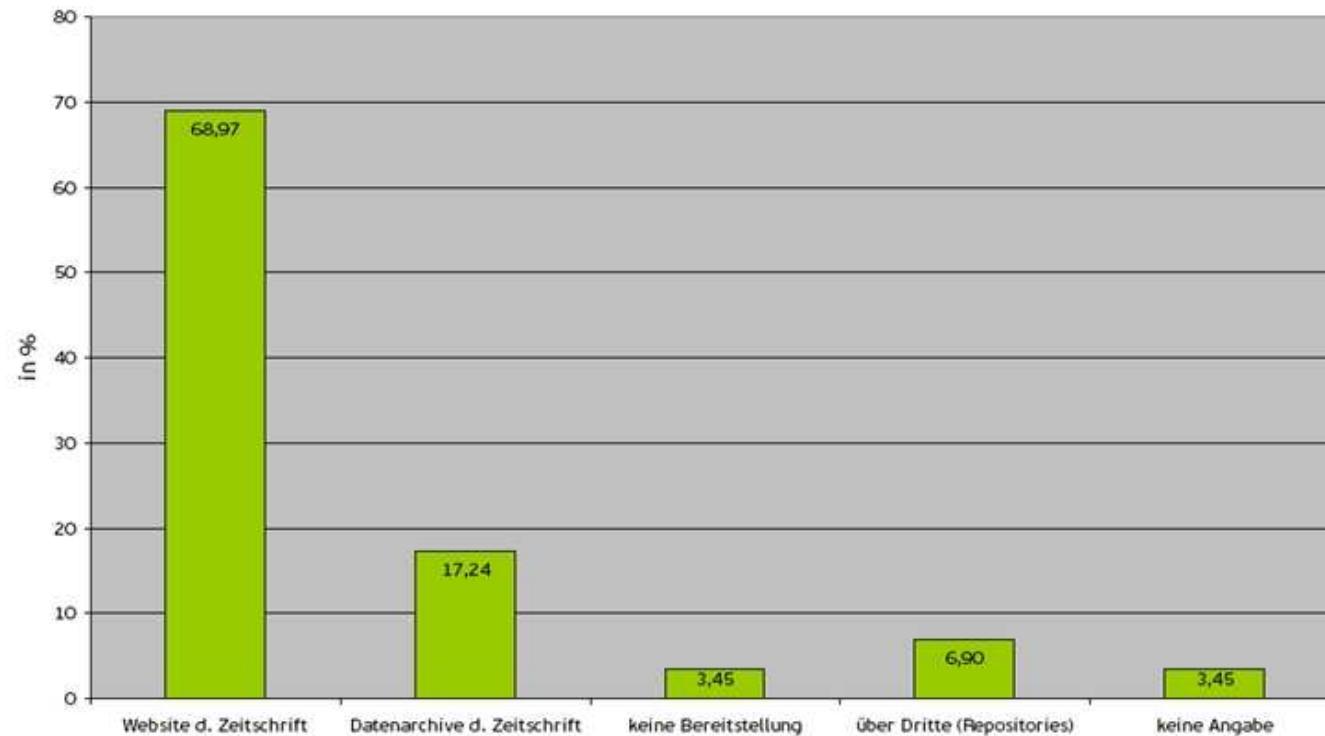
➤ Fast 90% der Richtlinien erfordern eine Einreichung vor der Publikation des Artikels

## Welche Daten müssen Autoren einreichen?



> Vor allem Datensätze werden eingefordert. Beschreibungen und Programme fordern nur noch  $\frac{2}{3}$  der Zeitschriften, den Berechnungscode nur noch knapp > 50%.

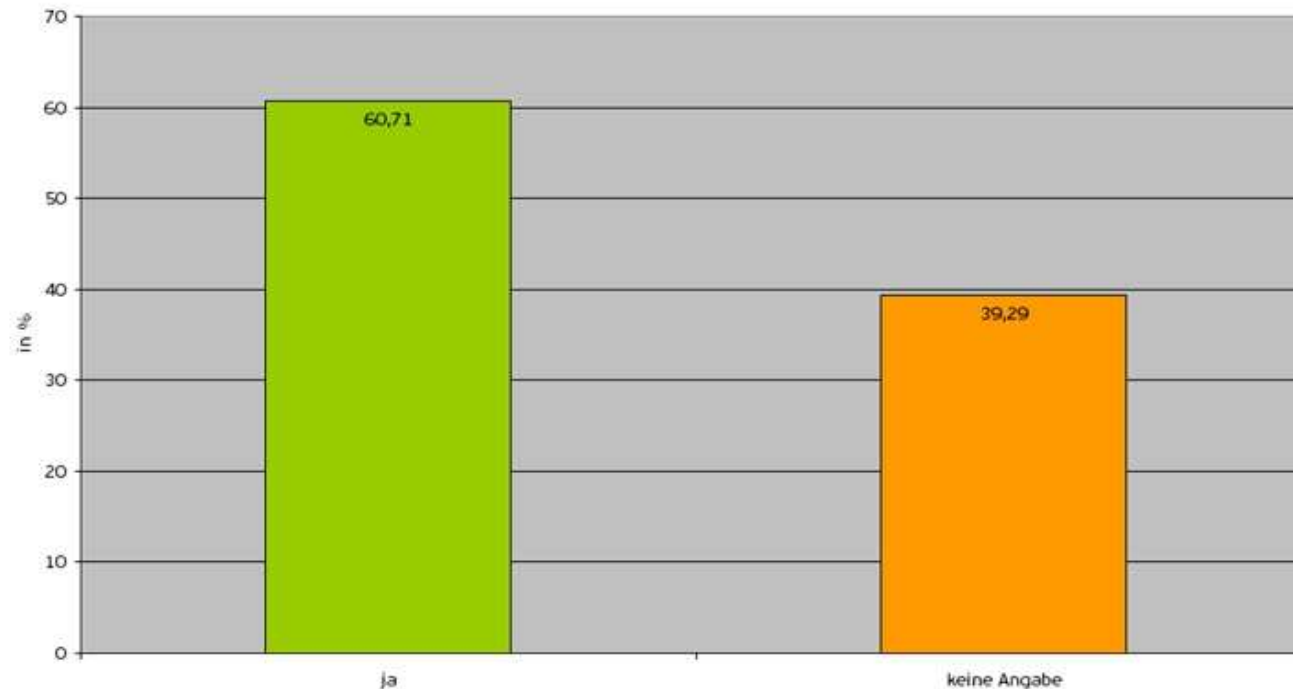
## Wie werden die Daten bereit gestellt?



➤ Die Bereitstellung v. Forschungsdaten in Fachzeitschriften erfolgt vor allem über „Supplementary Materials“ der Verlagswebseiten.

# Vorgaben bei Ausnahmen von der Policy?

Gibt es definiertes Vorgehen bei d. Verwendung proprietärer Daten?  
(n=29)



> > 60% der Richtlinien haben ein definiertes Vorgehen bei Nutzung propriet. Daten



# Ergebnisse der Studie

- > Anteil von Zeitschriften mit Data Availability Policy ist noch gering – Steigerungen sind festzustellen
- > Richtlinien unterscheiden sich in Umfang und Qualität erheblich
- > Mehrheit der Richtlinien ist verpflichtend
- > Fast alle Zeitschriften fordern Übermittlung der Daten vor der Publikation eines Artikels
- > Weniger als  $\frac{2}{3}$  aller Zeitschriften fordern Übermittlung von Programmen und Beschreibungen.
- > Noch starke inhaltliche Defizite bei 50% der Data Policies (-> Übermittlung von Code)
- >  $>\frac{3}{4}$  der Richtlinien ermöglichen Ausnahmen v. der Policy – aber nur 60% haben diesbezüglich ein Vorgehen definiert.
- > Bereitstellung von publikationsbezogenen Forschungsdaten ist stark verbesserungswürdig

# Fazit

Was benötigen wir, um Replikationen in den Wirtschaftswissenschaften zu vereinfachen?

Gefördert durch:

**DFG**

**EDaWaX**  
European Data Watch Extended

**RatSWD.**  
Rat für Sozial- und  
WirtschaftsDaten

**ZBW** Leibniz-Informationszentrum  
Wirtschaft  
Leibniz Information Centre  
for Economics

## Gap - Analyse

- > Anreize für Forscher „ihre“ Daten bereit zu stellen, werden benötigt
- > Mehr Fachzeitschriften denen das Thema „Replikation“ wichtig ist (-> Data Availability Policies)
- > Standards für Replikationsanforderungen (= zu übermittelnde Daten)
- > Infrastrukturelle Lösungen zur Übermittlung, zum Auffinden und Nachnutzen von Forschungsdaten – das betrifft Storage und Hosting, aber auch Metadatengenerierung und Retrieval Dienste.

**Vielen Dank für Ihre Aufmerksamkeit!**

Fragen?

Kontakt: [s.vlaeminck@zbw.eu](mailto:s.vlaeminck@zbw.eu) | [www.edawax.de](http://www.edawax.de)