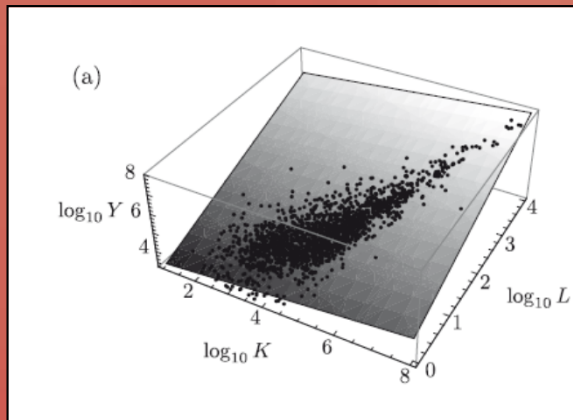# 'Work Closely with the Community!'
Experiences from a Project on Research Data Management for Economic Journals

Source: Economics-EJournal

Source: American Economic Review

Sven Vlaeminck
ZBW – German National Library of Economics / Leibniz-Information Centre for Economics
LIBER Conference | 1st of July 2016 | Helsingfors | Suomi

Funded by
DFG

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Agenda:

> 1. Background and Aims of EDaWaX
> 2. The Project's Approach
>   – 2.1 Analysis
>   – 2.2 Conception
>   – 2.3 Technical Implementation
>   – 2.4 Dissemination & Community Feedback
> 3. Lessons learned

EDaWaX
European Data Watch Extended

Max Planck Institute for
Innovation and Competition

DIW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics
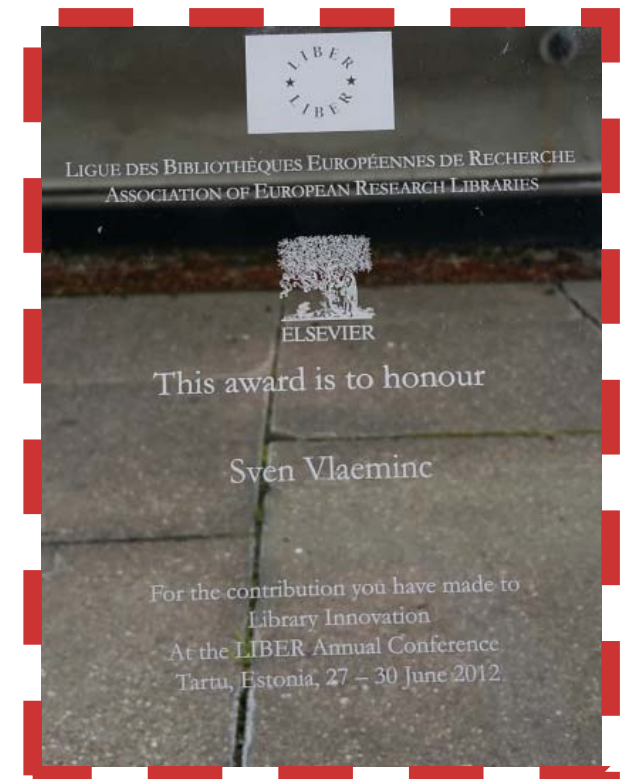
1. Introduction:
The Background of the EDaWaX-Project

> Changing patterns of published economic research and the replication crisis in the social sciences.

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

1.1 Changing Patterns of published Research

> Hamermesh analysed 728 articles published in economics top-journals between 1963 and 2011.

> Results: Methodology used changed fundamentally.

– Research based on borrowed or self-collected datasets, experimental designs or simulations reached 80% in 2011.

– In 1963 this share was below 50%.*

> More than ever, research in economics is ‚data-based'

*Hamermesh, D. S. (2012), Six Decades of Top Economics Publishing: Who and How?
National Bureau of Economic Research, Working Paper 18635

1.2 The Replication Crisis in Economic Sciences

> Dewald et al. (1986) tried to replicate 54 articles of the *Journal of Money, Credit, and Banking (JMCB)*.

  – They succeeded two times (3.7%).

> Hubbard & Armstrong (1994) analysed 1,120 papers in marketing journals.

  – <u>None</u> of these were replications; 1.8% were extensions.

  – Of these extensions, 60% found conflicting results

> McCullough et al. (2008) tried to replicate 117 articles of the *Federal Reserve Bank of St. Louis Review*.

  – They were able to replicate 9 articles (7.7%).

> Chang & Li (2015) successfully replicated 29 of 59 papers (49%) from seven journals *with authors' support*.

➔ …most often, economic research is <u>not </u>replicable.

1.3 Aim and initial Position of the Project:

> Aim of the project was to develop & to implement a research data archive for journals in economic sciences.

> The project consortium consisted of scientific infrastructure providers, economists and editors of scholarly journals.

> Project was located in a field, where we had limited knowledge prior to launching the project. => further expertise needed!

2. The Approach of EDaWaX

From an exploration of a new field of work towards first steps into the research community.

(Funding phases: 2011-2013 and 2014 - 2016)

EDaWaX
European Data Watch Extended

Max Planck Institute for
Innovation and Competition

DIW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Project phases:

Phase 1
ANALYSIS

- Evaluation of incentives to „share" research data
- Analysis of data policies of journals in business studies.
- Categorisation of research data used for creating economic articles (e.g. legal opportunities to „share" research data)

Phase 2
Conception

- Findings of analysis-phase lead to compilation of specifications to develop the application.
- Descision for a metadata schema.
- Evaluation of possible software solutions.

Phase 3
Technical Development

- Technological enhancement of EDaWaX's application for journals' data archives:
  * automated DOI-assignment,
  * metadata dissemination,
  * integration of data catalogues.
- Building up a sustainable infra-structure framework.

Phase 4
Dissemination

- Publication of project outputs
- Expansion of our user community
- Roadshows
- Workshops for editors of journals in economic sciences.

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW

ratSWD.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Project phases:

Phase 1
ANALYSIS

- Evaluation of incentives to „share" research data
- Analysis of data policies of journals in business studies.
- Categorisation of research data used for creating economic articles (e.g. legal opportunities to „share" research data)

Phase 2
Conception

- Findings of analysis-phase lead to compilation of specifications to develop the application.
- Descision for a metadata schema.
- Evaluation of possible software solutions.

Phase 3
Technical Development

- Technological enhancement of EDaWaX's application for journals' data archives:
  * automated DOI-assignment,
  * metadata dissemination,
  * integration of data catalogues.
- Building up a sustainable infra-structure framework.

Phase 4
Dissemination

- Publication of project outputs
- Expansion of our user community
- Roadshows
- Workshops for editors of journals in economic sciences.

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW

atSWD.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Step 1: Analysing the field

> First of all, we had to collect some information on the status quo in economic sciences:

- How many journals have data policies, which include replication files into the review and publishing process?
- How are these policies shaped? Do these policies forster replicability? Do journals enforce their data policies?
- How do journals with data policies provide replication data to would-be replicators?
- Do researchers in economic sciences share their data? Which incentives exist for researchers to share their data?
- Are there any infrastructure providers that offer services for journals in economics and business studies?
- …

Some Findings of our Analyses:

> ~15% of 346 journals evaluated have a *data availability policy (DAP)*.

> 6% have an old fashioned „policy" that relies on author's willingness to participate in the replication of his/her results (does not work!).

> Most of the data policies are robust, but some are merely window dressing.

> Only 48% of 259 articles published by 19 journals with a DAP honoured the data policy of the particular journal.

> ~ 84% of all journals with DAP provide replication files by publisher's website (often zip-files attached to the article) -> bad practise!.

> Data Sharing among economists is rare. Only 2.5% of 488 researchers surveyed regularly shared replication files.

> In 2011/2012, almost no suitable services (-> „replication server") for journals were available. But: In the last years, more and more RDCs implemented such services.

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Project phases:

Phase 1
ANALYSIS

- Evaluation of incentives to „share" research data
- Analysis of data policies of journals in business studies.
- Categorisation of research data used for creating economic articles (e.g. legal opportunities to „share" research data)

Phase 2
Conception

- Findings of analysis-phase lead to compilation of specifications to develop the application.
- Descision for a metadata schema.
- Evaluation of possible software solutions.

Phase 3
Technical Development

- Technological enhancement of EDaWaX's application for journals' data archives:
  * automated DOI-assignment,
  * metadata dissemination,
  * integration of data catalogues.
- Building up a sustainable infra-structure framework.

Phase 4
Dissemination

- Publication of project outputs
- Expansion of our user community
- Roadshows
- Workshops for editors of journals in economic sciences.

EDaWaX
European Data Watch Extended

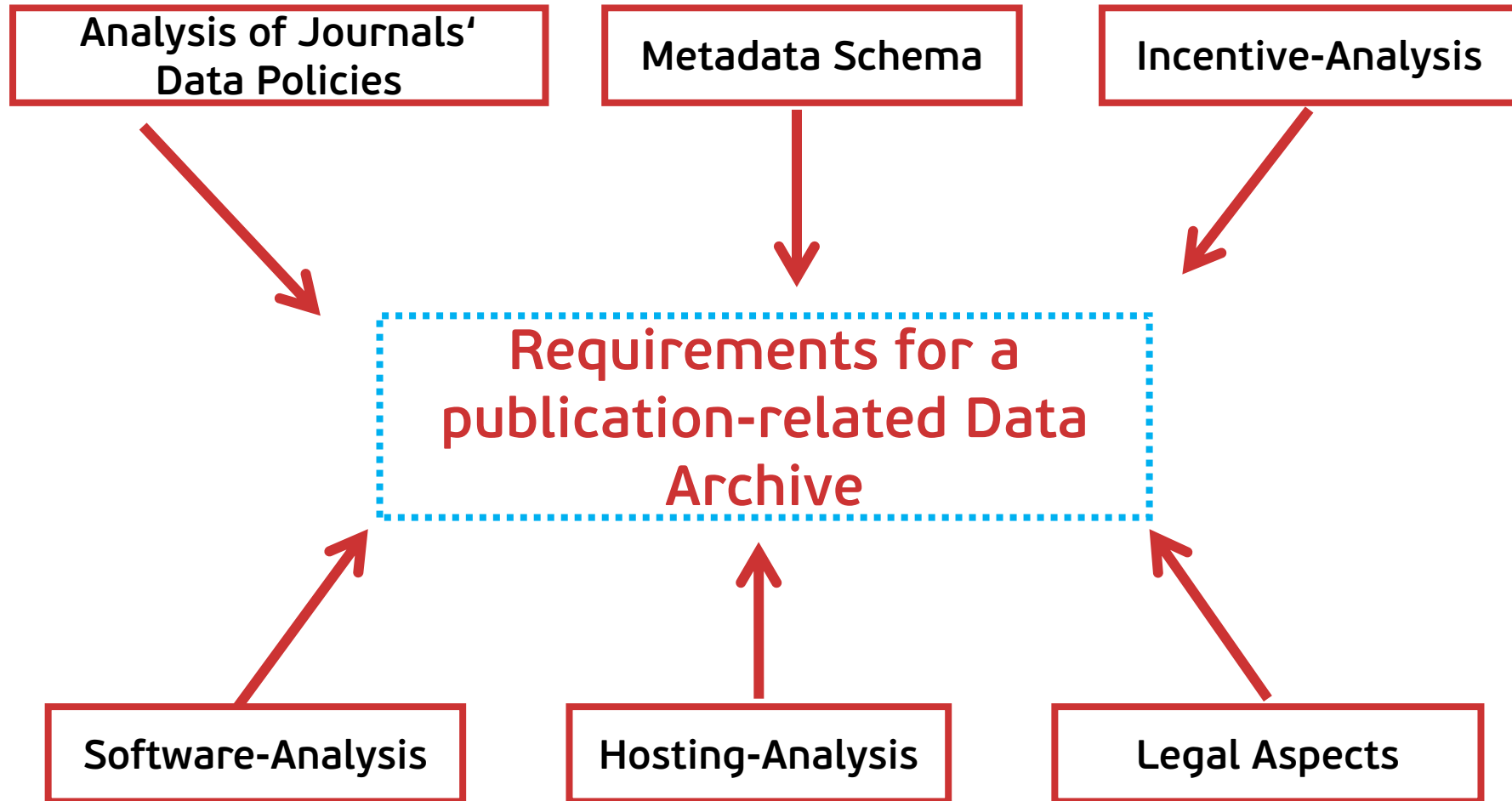Max Planck Institute for Innovation and Competition

DIW

SWD.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Step 2: Conception of the Software based on the Outcomes of our Analyses

Analysis of Journals' Data Policies

Metadata Schema

Incentive-Analysis

Requirements for a publication-related Data Archive

Software-Analysis

Hosting-Analysis

Legal Aspects

Step 2: Conception of the Software based on the Outcomes of our Analyses

> Next, we analysed and rated potential software solutions for our future data archive.

> Based on our specification sheet, we chose CKAN – a software, that is broadly used in the context of open data portals (www.ckan.org)

Project phases:

Phase 1
ANALYSIS

- Evaluation of incentives to „share" research data
- Analysis of data policies of journals in business studies.
- Categorisation of research data used for creating economic articles (e.g. legal opportunities to „share" research data)

Phase 2
Conception

- Findings of analysis-phase lead to compilation of specifications to develop the application.
- Descision for a metadata schema.
- Evaluation of possible software solutions.

Phase 4
Dissemination

- Publication of project outputs
- Expansion of our user community
- Roadshows
- Workshops for editors of journals in economic sciences.

Phase 3
Technical Development

- Technological enhancement of EDaWaX's application for journals' data archives:
  * automated DOI-assignment,
  * metadata dissemination,
  * integration of data catalogues.
- Building up a sustainable infra-structure framework.

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW

atSWD.

ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Step 3: Software's technical Development

> First, we integrated some basic functionalities (-> Pilot application).

> Expansion of features and advanced functionalities on a step-by-step basis.

  - Adjustment of metadata fields for different ‚resources'
  - Integration of authority files (GND via Lookup-Service)
  - Semi-automated DOI-assignment (by da|ra via API)
  - Optimisation of resources' metadata to facilitate integration of research data in disciplinary portals and in Google Scholar.

> Since end of 2015 we prepared the implementation of a permanent service (ZBW Journal Data Archive) at the library.

Datasets   Journals   About   Search

Welcome to the ZBW Journal Data Archive!

The journal data archive offers a possibility for editors of scholarly journals in economics and business studies to facilitate the management of replication data.

Authors of these journals are invited to submit their replication files to the data archive. Preliminary, a documentation on how to use the application is available here. A documentation in German is available here.

The ZBW Journal Data Archive is a service for journal editors and successor of a pilot application that has been developed by the DFG-funded project EDaWaX.
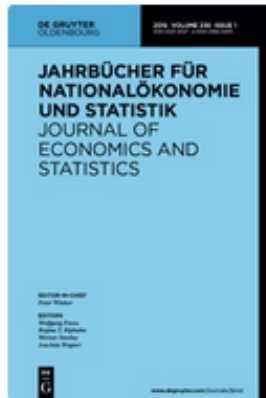
Search data

E.g. environment

Popular tags   Germany   economics
youth unemployment

EDaWaX
European Data Watch Extended

Max Planck Institute for
Innovation and Competition

DIW SOEP RatSWD. ZBW   Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Still Different After All These Years (replication data)

Joachim Wagner

Dataset | Activity Stream | Related

Still Different After All These Years (replication data)

Notes on the firm level data used in the paper "Still different after all these years - Extensive and intensive margins of exports in East and West German manufacturing enterprises" by Joachim Wagner

The firm level data used in the paper are taken from surveys that are administered by the statistical offices. These data are strictly confidential, but they can be used by researchers inside the research data centers of the statistical offices. For more information on how to access the data see www.forschungsdatenzentrum.de and go to the link . The computations for this paper use a tailor-made data set that combines firm level data from various surveys; for details see section 2 of the paper and the literture cited there.

The Stata-do-file "east-west-01.do" prepares the data and computes descriptive statistics and tests for differences between east and west German firms. The Stata-do-file "east-west-06.do" estimates the empirical models for the export margins and performs the decomposition analyses.

Researchers interested in using the data and do files for replication and extensions are encouraged to contact the author vial email at .

Data and Resources

Notes on the firm level data used.txt
Notes on data sets and code — Explore

east-west-01.do
The Stata-do-file east-west-01.do prepares the data and computes descriptive... — Explore

east-west-06.do
The Stata-do-file east-west-06.do estimates the empirical models for the... — Explore

Citation

Wagner, Joachim (2016): Still Different After All These Years (replication data). Version: 1. Journal of Economics and Statistics. Dataset. http://dx.doi.org/10.15456/jbnst.2016117.070159

Related Publication

doi:10.1515/jbnst-2015-1011

Journal of Economics and Statistics

The Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik) is a scientific journal published in Germany since 1863, since 2016 by DeGruyter. The... read more

Social
Google+
Twitter
Facebook

JAHRBÜCHER FÜR NATIONALÖKONOMIE UND STATISTIK
JOURNAL OF ECONOMICS AND STATISTICS

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW SOEP   RatSWD.   ZBW   Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Project phases:

Phase 1
ANALYSIS

- Evaluation of incentives to „share" research data
- Analysis of data policies of journals in business studies.
- Categorisation of research data used for creating economic articles (e.g. legal opportunities to „share" research data)

Phase 2
Conception

- Findings of analysis-phase lead to compilation of specifications to develop the application.
- Descision for a metadata schema.
- Evaluation of possible software solutions.

Phase 3
Technical Development

- Technological enhancement of EDaWaX's application for journals' data archives:
  * automated DOI-assignment,
  * metadata dissemination,
  * integration of data catalogues.
- Building up a sustainable infra-structure framework.

Phase 4
Dissemination

- Publication of project outputs
- Expansion of our user community
- Roadshows
- Workshops for editors of journals in economic sciences.

EDaWaX
European Data Watch Extended

Max Planck Institute for Innovation and Competition

DIW

SWD.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Step 4: Dissemination & Community Feedback

> Since late 2011, we announced the aims of the project in the economic community:

- Part of it was a blog, where we published all types of news with relation to reproducible research (in total ~140 posts)
- Research articles published by the project attracted attention in the community (-> networks, community building).
- Workshops and panel discussions at annual meetings of learned societies helped to stay in touch with researchers.
- Special workshops for authors and editors were extremely valuable to develop a software in accordance with journals' and researchers' needs.

Community-Feedback:

Panel discussion on „data sharing" at the annual meeting of the German learned society for economics. (06.09.2013)

Evaluation workshop with editors of social sciences' journals (18.11.2013)

IW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

3. Lessons learned

EDaWaX
European Data Watch Extended

Max Planck Institute for
Innovation and Competition

DIW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

A few ‚take home messages':

> Involve researchers in the project consortium.
  - Offers subject-specific perspectives (very important!)
  - Facilitates contacts to the other researchers in the field
> For developing software, periodical workshops with the intended target group are a crucial success factor.
  - These workshops offer useful feedback for the IT-work.
  - It becomes more likely, that researchers really use such tools.
> Own research offers more visibility and contacts.
  - Joint working on eye level with researchers becomes more likely.

Some more ‚take home messages'

> Don't wait for the researchers to get in touch with you.
  - …because usually, they won't!
  - Be proactive, ask for infrastructural needs, talk about your services.
> Relations to subject communities need time and patience.
  - …but it is worth the efforts…
  - …and there is no alternative! Who else should use your super-duper IT-infrastructure if not the intended target group?

Further reading/linking:

> ZBW's Journal Data Archive : journaldata.zbw.eu.

> Publications, presentations (and also replication files of our research papers):
  – http://www.edawax.de/downloads/ and
  – http://www.edawax.de/downloads/project-output-funding-phase-i/

> The program code of our CKAN application is available on Github:
  – https://github.com/hbunke

Thank you
very much for your attention!

...do you have questions or comments?

Sven Vlaeminck

ZBW – Leibniz-Information Centre for Economics

Hamburg / Germany

Tel.: +49 40 – 42834 419

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

EDaWaX
European Data Watch Extended

Max Planck Institute for
Innovation and Competition

DIW SOEP RatSWD. ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics