



Leibniz Institute
for the Social Sciences

Access to social science research data by an open API

IASSIST 2017 in Lawrence, Kansas:

Data in the Middle: The common language of research

May 23-26, 2017

Wolfgang Zenk-Möltgen, Reiner Mauer

Outline

- Social science research data
- Current services at GESIS
- Format conversions
- Provide access
- Build an easy to use API
 - ▶ RESTful
 - ▶ Documented by OpenAPI
- Future use

Social science research data

- At the GESIS Data Archive ...
 - ▶ the social science research data have a focus on sociology and political science
 - ▶ the data are mainly quantitative rectangular by structure
 - ▶ predominant data collection method is survey
- Current expansions are ...
 - ▶ social media and internet data
 - ▶ experimental data
 - ▶ georeferenced survey data

Social science research data

- A variety of nations



Social science research data

- A variety of topics



Social science research data

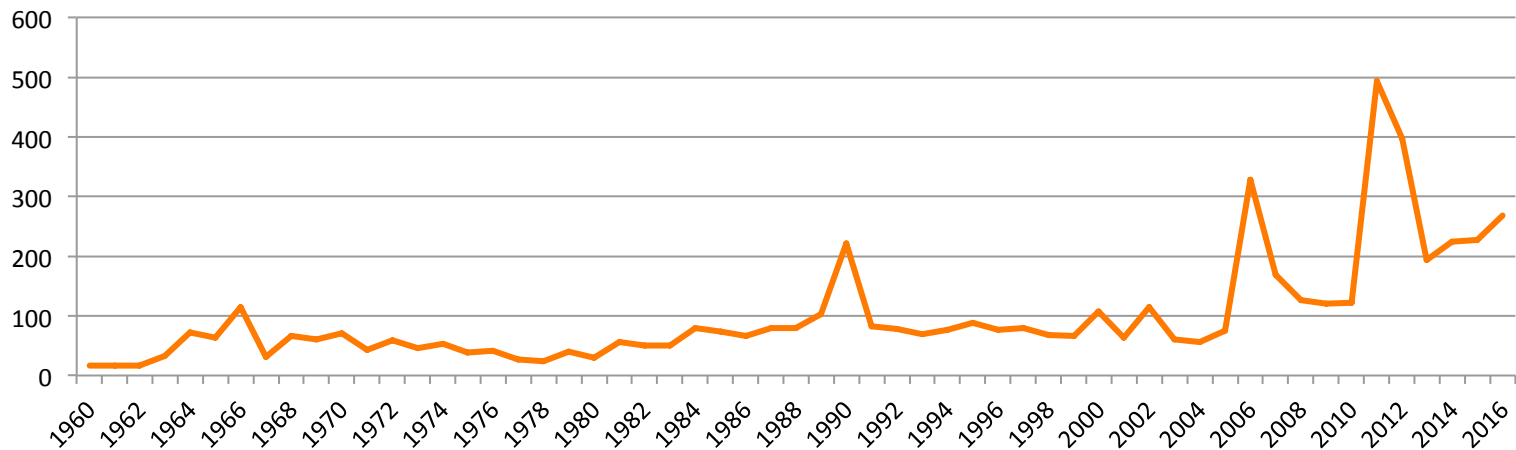
■ A variety of institutions



Social science research data

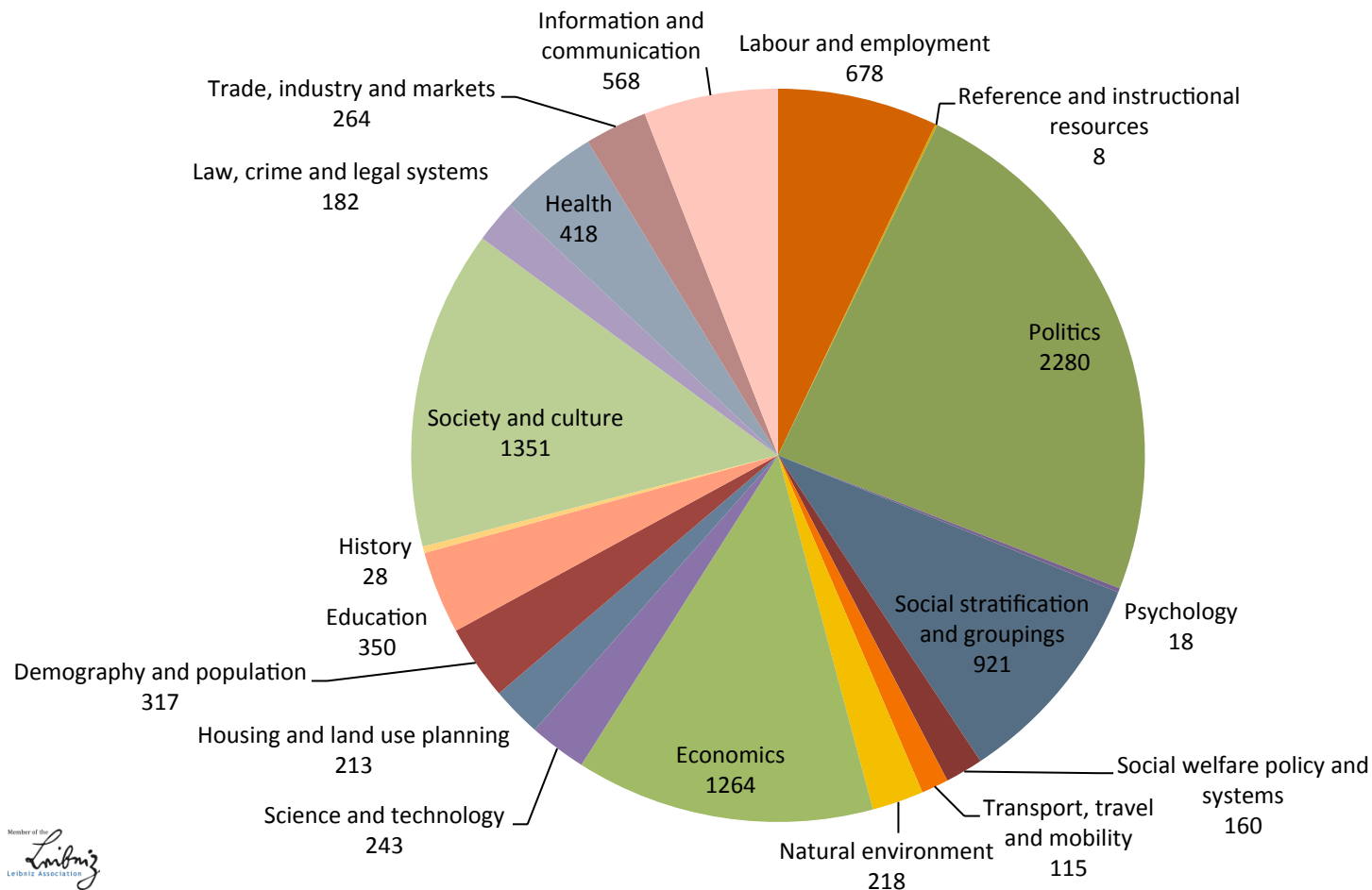
- Increase of published studies per year

Studies by publication year



Social science research data

- 5767 studies by different topic areas



Current services at GESIS

- How do we provide access to data?
- The Data Catalogue DBK
 - ▶ Download data sets
 - ▶ Order data sets
- datorium
 - ▶ Download data sets
- HISTAT
 - ▶ Download data

Current services at GESIS: DBK

gesis

Leibniz-Institut
für Sozialwissenschaften

Deutsch

Contact

DBK Home

Search

Browse

Overview

Account

News

wolfgang.zenk-moeltgen@gesis.org

GN0nnn (158)

1

2

3

4

5

6









7

8


9

10

>>

No.	Group number	Group Title	Order
1	GN0001	International Social Survey Programme (ISSP)	
2	GN0002	EB - Flash Eurobarometer	
3	GN0003	Travel Surveys	
4	GN0004	Time Budget Study	
5	GN0005	EB - Central and Eastern Eurobarometer	
6	GN0006	EB - Candidate Countries Eurobarometer	
7	GN0007	ALLBUS	
8	GN0008	EB - Standard and Special Eurobarometer	
9	GN0009	European Values Study (EVS)	

Current services at GESIS: DBK




**Leibniz-Institut
für Sozialwissenschaften**


[Deutsch](#) [Contact](#)

DBK Home ▾
Search ▾
Browse ▾
Overview ▾
Account ▾
News
wolfgang.zenk-moeltgen@gesis.org

ZA5760: Intermediate inquiry 2016 of Long-term Panel started in 2005,2009 and 2013 (GLES)

Bibliographic Citation	Methodology	Data & Documents	Errata & Versions	Further Remarks
<div style="display: flex; justify-content: space-between; align-items: center;"> Publications Groups </div> <div style="margin-top: 10px;"> <p>Citation ⓘ </p> <p><i>Roßteutscher, Sigrid; Schmitt-Beck, Rüdiger; Schoen, Harald; Weißels, Bernhard; Wolf, Christof; Preißinger, Maria; Rudnik, Agatha; Wuttke, Alexander (2017): Intermediate inquiry 2016 of Long-term Panel started in 2005,2009 and 2013 (GLES). GESIS Data Archive, Cologne. ZA5760 Data file Version 1.0.0, doi:10.4232/1.12750</i></p> </div>				
Study No.	ZA5760			
Title	Intermediate inquiry 2016 of Long-term Panel started in 2005,2009 and 2013 (GLES)			
Current Version	1.0.0, 2017-2-24, doi:10.4232/1.12750			
Date of Collection	05.10.2016 - 15.12.2016			
Principal Investigator/ Authoring Entity, Institution	<ul style="list-style-type: none"> • Roßteutscher, Sigrid - Universität Frankfurt • Schmitt-Beck, Rüdiger - Universität Mannheim • Schoen, Harald - Universität Mannheim • Weißels, Bernhard - Wissenschaftszentrum Berlin für Sozialforschung 			



Current services at GESIS: DBK



Leibniz-Institut
für Sozialwissenschaften


[Deutsch](#)
[Contact](#)

DBK Home ▼ Search ▼ Browse ▼ Overview ▼ Account ▼ News
wolfgang.zenk-moeltgen@gesis.org

ZA5760: Intermediate inquiry 2016 of Long-term Panel started in 2005,2009 and 2013 (GLES)

Bibliographic Citation Publications Groups	Methodology	Data & Documents	Errata & Versions	Further Remarks
Dataset	Number of Units: 1848			
	Number of Variables: 169			
	Analysis System(s): SPSS, Stata			
Availability ⓘ	A - Data and documents are released for academic research and teaching.			
Download of Data and Documents ⓘ	All downloads from this catalogue are free of charge. Data-sets available under access categories B and C must be ordered via the shopping cart. Charges apply! Please respect our Terms of use .			




Datasets

Questionnaires

Other Documents

DDI Documents



- [ZA5760_missing.do](#) definition of missing values (for Stata) (Dataset) 2 KBytes
- [ZA5760_open-ended_v1-0-0.csv](#) (Dataset) 117 KBytes
- [ZA5760_v1-0-0.dta](#) (Dataset STATA) 1 MByte

Current services at GESIS: DBK

- DBK data file download and orders

gesis Leibniz-Institut für Sozialwissenschaften

Deutsch Contact

DBK Home Search Browse Overview Account News wolfgang.zenk-moeltgen@gesis.org

Data Catalogue Download

Close page

Download 'ZA5760_v1-0-0.dta' (Dataset)

Please enter the purpose of your data usage:

I use the data Please enter the purpose of your usage ▼

☒ With this download I agree to the **'terms of use'** of GESIS.

Download

* Only available for access class 0. For studies of access classes A, B, C: Data are not free to be downloaded, please contact our Data Service: [datenservice.das \(at\) gesis.org](mailto:datenservice.das@gesis.org).

Current services at GESIS: DBK

gesis
Leibniz-Institut
für Sozialwissenschaften
Deutsch Contact

DBK Home ▾ Search ▾ Browse ▾ Overview ▾ Account ▾ News
not logged on

Data Catalogue Shopping Cart Order

Charges apply for data-sets ordered via this shopping cart!

Your order will be processed as soon as possible by the data service team during normal working hours. Data and documentation will be supplied via ftp-delivery or on data storage medium via normal mail.

The following products and studies are in your shopping cart		
Product	Product name	Count, Purpose
Study	Study Title	Purpose
ZA7474	Consciousness of Nature 2015	<div>Please enter the purpose of your usage ▾</div> <div style="text-align: right;"> delete </div>
Group	Group Name	Purpose
desired data format (if available):		SPSS-Windows ▾
desired delivery (if available):		FTP ▾
Please enter the purpose of your data usage:		<input type="checkbox"/> different purposes for list items
I use the data		

Please enter the purpose of your usage ▾

Current services at GESIS: datorium

gesis

datorium

[Diese Seite auf Deutsch](#) | [Login](#) | [Register](#) | [Contact](#)

Search

Add data

About datorium ▶

Mitglied der

Welcome to datorium

datorium is a data repository service for the social science and economic science research community. It provides a user-friendly tool for the autonomous documentation, upload and publication of research data. This includes survey data, administrative data, tabular data and outputs of statistical analyses. Syntax files may also be submitted.

Enabling replications
Furthermore, datorium offers journals a service to support replicability of empirical research by sharing research data of published journal articles. For more information visit www.replikationsserver.de

Documentation and Upload

Storage and Review

Publication

Search and Download

Search:

New Data

House Building and Retail Sales in the German Democratic Republic 1952 to 1989	2017-05-15
German Index of Socioeconomic Deprivation (GISD) Version 1.0	2017-05-12
Code: Einkommensunterschiede in der Mortalität in Deutschland – Ein empirischer Erklärungsversuch	2017-05-03

Contact

datorium Team
GESIS Data Archive for the Social Sciences
Tel.: +49(221)-47694-417
 E-Mail

Current services at GESIS: datorium

Diese Seite auf Deutsch | Login | Register | Contact

Filter Publication Year 2017 (13) 2016 (33) 2015 (23) 2014 (12) 2013 (3) 2012 (1) 2011 (1) Availability Subject Area Topic Classification Geographical Area Data Collection Method Replication Server	<div style="background-color: #e0e0e0; padding: 5px;">Search</div> <div style="margin-top: 10px;"> <input style="width: 100%;" type="text"/> <input style="margin-left: 10px;" type="button" value="Go"/> </div> <div style="background-color: #e0e0e0; padding: 5px;">Sort options</div> <div style="background-color: #e0e0e0; padding: 5px; margin-top: 10px;">Results</div> <p>Now showing items 1-10 of 86</p> <div style="text-align: center; margin-bottom: 10px;"> 1 2 3 4 ... 9 > </div> <div style="background-color: #e0e0e0; padding: 5px; margin-bottom: 10px;"> House Building and Retail Sales in the German Democratic Republic 1952 to 1989 Primary Researcher / Institution: Driesch, Ellen von den; WZB Berlin Social Science Center Publication Year: 2017 Availability: Freier Zugang (mit Registrierung) </div> <div style="background-color: #e0e0e0; padding: 5px; margin-bottom: 10px;"> German Index of Socioeconomic Deprivation (GISD) Version 1.0 Primary Researcher / Institution: Kroll, Lars Eric; Robert Koch-Institut Publication Year: 2017 Availability: Freier Zugang (ohne Registrierung) </div> <div style="background-color: #e0e0e0; padding: 5px; margin-bottom: 10px;"> Code: Einkommensunterschiede in der Mortalität in Deutschland – Ein empirischer Erklärungsversuch Primary Researcher / Institution: Kröger, Hannes; Sozio-oekonomisches Panel am DIW Berlin Publication Year: 2017 Availability: </div>
--	---

Current services at GESIS: datorium

Diese Seite auf Deutsch | Login | Register | Contact

gesis

datorium

[Home](#)

[Home](#)

[View Item](#)

[Help](#)

Search

Add data

About datorium ▶

General Description [Back](#)

cite this data

Title	Publication Bias in the German Social Sciences
URI	http://dx.doi.org/10.7802/73
Primary Researcher	Berning, Carl; Department of Political Science, University of Mainz Weiß, Bernd; Institute of Sociology and Social Psychology, University of Cologne
Publication Year	2015
Availability	Restricted Access
Subject Area	Basic Research in the Social Sciences
Abstract	Systematic research reviews have become essential in all empirical sciences. However, the validity of research syntheses is threatened if the preparation, submission or publication of research findings depends on the statistical significance of these findings. The present study investigates publication bias in three top-tier journals in the German social sciences, utilizing the caliper test. For the period between 2001 and 2010, we have collected 156 articles that appeared in the Kölner Zeitschrift für Soziologie und Sozialpsychologie (KZfSS), the Zeitschrift für Soziologie (ZfS) and the Politische Vierteljahresschrift (PVS). In all three journals, we found empirical evidence for the existence of a publication bias at the 10% level. We also investigated possible causes linked to this bias, including single versus multiple authorship as well as academic degree. We find only weak support for the relationships between individual author characteristics and publication bias.
Publications	Berning, C., & Weiß, B. (2015). Publication Bias in the German Social Sciences: An Application of the Caliper Test to Three Top-Tier German Social Science Journals. Quality & Quantity.; http://dx.doi.org/10.1007/s11135-015-0182-4

Mitglied der


Leibniz-Gemeinschaft

Files in this submission


Files	Size	Format	Download Option
berning.weiss.2015-01.zip	486.5Kb	application/zip	for scientific research (incl. PhD) ▼
			Download


🔗 **File checksum:** MD5:863cf263a7b416661dde94f002de1608

Current services at GESIS: HISTAT

gesis 

histat: 7.640.310 Values 367.234 Time series
Historical statistics

Home Data About Login 

Searchtext 

Historical Time Series



Welcome to the online-database **histat**. The database has been offered via GESIS since 2004. A substantial revision has been made in 2012. The number of database calls has been more than doubled within the last two years. This year (in 2012) we will attain more than 3.000 registered users. In the meantime, the database encloses a quarter of a million time series, 359 studies, and more than five million values.

Citation

Kindly cite the sources of the used data and the respective publication of the downloaded studies in your own publication according to the scientific conventions. Recommendations regarding the citation of historical data can be found under the topic "[Bibliographic citation of research data and study related documents](#)".

In cooperation with

DI STATIS
Statistisches Bundesamt

What others say about histat

"A really good and important project."

Prof. Dr. Martin Schulze Wessel
Chairman des Verbandes der Historiker und
Historikerinnen Deutschlands

"An indispensable data-repository for German
economic and social history."

Prof. Dr. Jörg Baten
General Secretary der International Economic
History Association

"It is most certainly not possible to make it more
comfortable."

Dr. Oliver Volckart
London School of Economics

"It's great that GESIS collects this data and
provides it to researchers."

Robin Winkler
DPhil candidate in Economic History, Oxford
University

Further cooperating Partners

Current services at GESIS: HISTAT

gesis

histat:

Historical statistics

7.640.310 Values 367.234 Time series

Home Data About Login

Searchtext 🔍

Categories | Periods | Names | New | Top


On this page you get the opportunity to select studies via their thematic allocation. All studies have been assigned exactly on this subject, which is the focus of the study.

Arbeit | Bauen | Bevölkerung | Bildung | Einkommen | Energie | Geld | Gesundheit | Handel | Hanse | Industrie | Innovation | Konjunktur | Kriminalität | Kultur | Landwirtschaft | Preise | Sozialstaat | Staatsfinanzen | Städte | Umwelt | Unternehmen | Verbrauch | Verkehr | Versicherungen | VGR | Wahlen | Übergreifend | SIMon: Social Indicators Monitor 1950-2013 | Demonstrationsbeispiele


ZA 8245	Handel	Hoffmann, Walther G. (1965 [2006]), Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts. Kapitel: Die Produktion von Handel, Banken, Versicherungen und Gaststätten.	40 Time series (1850 - 1959) 5 Tables	Details...
ZA 8258	Handel	Hoffmann, Walther G. (1965 [2006]), Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts. Kapitel: Der Außenhandel, die Zahlungsbilanz.	139 Time series (1836 - 1960) 14 Tables	Details...
ZA 8107	Handel	Dessauer, M. (1982 [2007]), Entwicklungstendenzen der betrieblichen Exportwirtschaft in Deutschland seit der Mitte des 19. Jahrhunderts unter besonderer Berücksichtigung der Exportorganisation	39 Time series (1872 - 1979) 8 Tables	Details...
ZA 8315	Handel	Buchheim, Christoph (1982 [2008]), Deutsche Gewerbeexporte nach England in der zweiten Hälfte des neunzehnten Jahrhunderts, 1854-1913.	103 Time series (1854 - 1913) 12 Tables	Details...
ZA 8313	Handel	Wagenführ, Rolf (1936 [2008]), Die Exportquote der deutschen Industrie von 1872 bis 1935.	111 Time series (1872 - 1935) 15 Tables	Details...
ZA 8363	Handel	Wulf, Jürgen (1968 [2009]), Der Deutsche Aussenhandel 1850 bis 1965. Entwicklung, Strukturwandlungen und Beziehungen zum Wirtschaftswachstum	226 Time series (1800 - 1966) 31 Tables	Details...


Current services at GESIS: HISTAT

- HISTAT time series data in tables per study

gesis 

hstat: Historical statistics 7.640.310 Values 367.234 Time series

Home Data About Login 










Searchtext 

Categories | Periods | Names | New | Top

Tableview

ZA 8245	Handel	Hoffmann, Walther G. (1965 [2006]), Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts. Kapitel: Die Produktion von Handel, Banken, Versicherungen und Gaststätten.	40 Time series (1850 - 1959) 5 Tables	Details...
---------	--------	--	---	------------

01. Die Produktion des Handels (1850-1959) (found timeseries: 17) Download

Produkte *	Nahrungsmittel	Brennstoffe	Holz	Baumaterial	Metallwaren	Papier	Chemikalien	Seife	Bier
Tables	Tab. 93	Tab. 93	Tab. 93	Tab. 93	Tab. 93	Tab. 93	Tab. 93	Tab. 93	Tab. 93
Sources	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...	Hoffmann, W. G. (196...
Notes	1913=100 bzw. 10000....	1913=100 bzw. 10000....	1913=100 bzw. 10000....	1913=100 bzw. 10000....	1913=100 bzw. 10000....	1913=100 bzw. 10000....	1913=100 bzw. 10000....	1913=100 bzw. 10000....	1913=100 bzw. 10000....
Graph									
1850	887	8	43	14	20				44
1851	864	9	43	15	21				44
1852	894	10	47	15	21				41
1853	913	11	47	17	23				42
1854	940	12	49	10	25				38
1855	962	14	48	12	27				38
1856	981	15	57	11	31				40
1857	1026	16	60	13	35				47
1858	1041	18	59	14	39				50
1859	1025	17	53	13	37				53
1860	1082	18	57	16	36				51
1861	1129	20	63	18	34				52
1862	1150	22	70	19	36				57
1863	1240	24	74	25	45				63
1864	1287	28	72	26	50				66

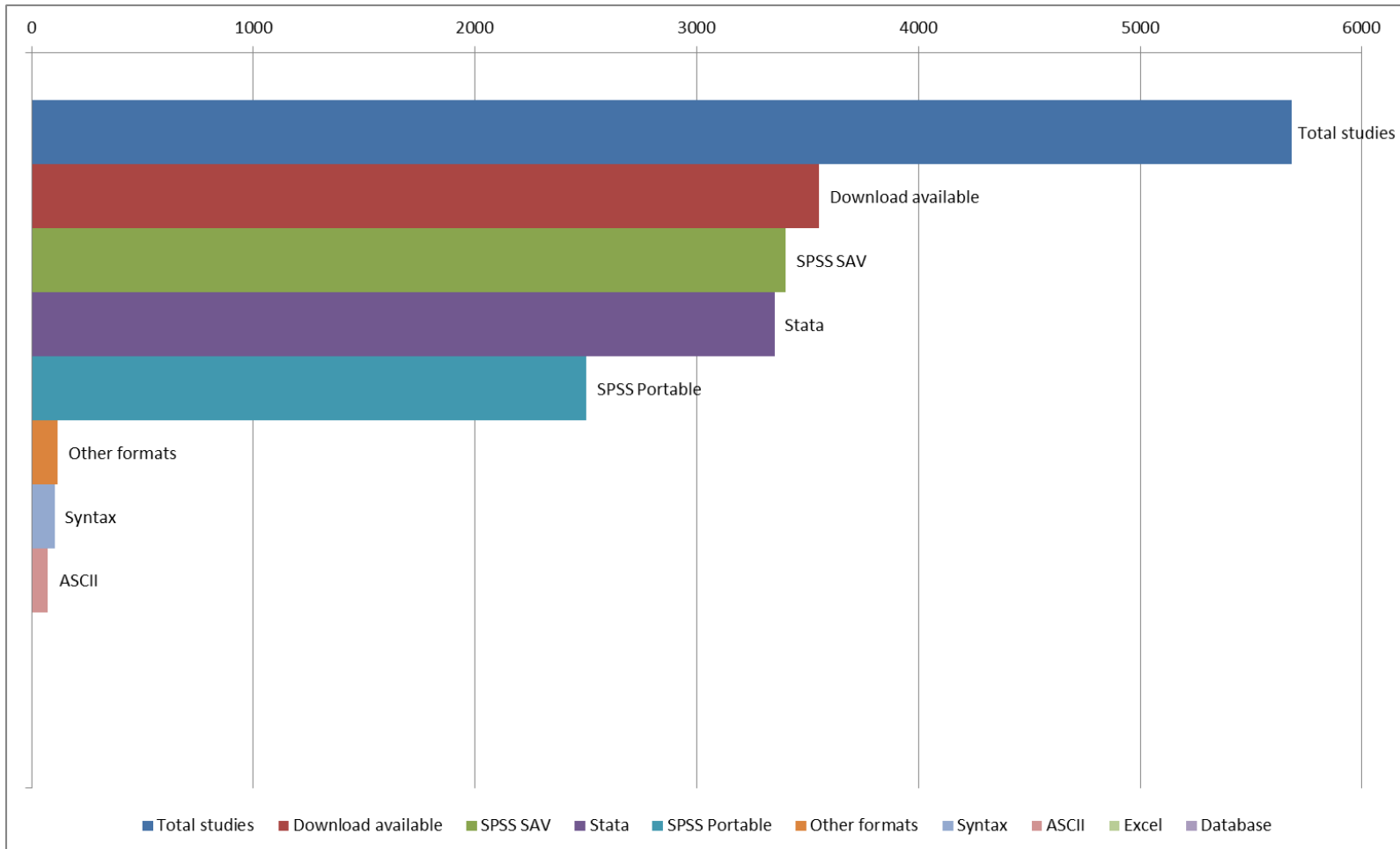
Format conversions

- Current file storage holds a variety of formats
- The Data Catalogue DBK
 - ▶ Predominantly SPSS (Portable, Windows)
 - ▶ STATA
 - ▶ Excel (CSV, XLS, XLSX)
 - ▶ ZIP-compressed files
 - ▶ Others
- datorium
 - ▶ What the researchers provide
- HISTAT
 - ▶ Excel (CSV, XLS, XLSX)

Format conversions

- DBK & HISTAT
 - ▶ internal long-term archiving repository
 - ▶ the originally submitted files
 - ▶ the processed files to be archived
 - ▶ the service files to be distributed
 - SPSS, STATA
 - Future: vendor independent, unicode based format
- Datorium
 - ▶ bistream preservation

Format conversions



Format conversions

- Tools to convert
 - ▶ PSPP
 - ▶ StatTransfer
- Issues with StatTransfer
 - ▶ ISO-8859-1 coded source files
 - ▶ Variable names in capital letters (sometimes)
 - ▶ Variable name replacement errors
 - ▶ Umlaut problems
 - ▶ Embedded newlines
 - ▶ Numbers shortened/rounded
 - ▶ minor updates already break compatibility

Format conversions

- Shell script with conversions and sanity checks
- Output: UTF-8 encoded, UNIX file format
 - ▶ <Study>[_<version>].csv (the data)
 - ▶ <Study>[_<version>].txt (metadata)
 - ▶ <Study>[_<version>].sps (SPSS syntax script, DOS)
 - ▶ <Study>[_<version>]-DDI-2_5.xml (metadata)
 - ▶ <Study>[_<version>].sql (SQL script)
- Special Thanks to *Andre Müller*

Format conversions

- Handle data

```
## Logging
echo "PSPP meta log:" > $logdir/${basename}.log_pspp
head -n `expr $startline - 1` pspp_meta.proc2 | grep -v "^$" >> $logdir/${basename}.log_pspp
head -n `expr $startline - 1` pspp_meta.proc2 | grep -v "^$"
echo -e "\nPSPP data log:" >> $logdir/${basename}.log_pspp
cat pspp_data.log >> $logdir/${basename}.log_pspp
# begin data file paranoia
rows=`grep "cases were transferred$" st_out-csv.stlog | cut -d\ -f1`
cols=`grep "^Input file has" st_out-csv.stlog | cut -d\ -f4`
h1=`awk -F"\t" '{print NF}' pspp_out.csv | uniq -c`
echo -n " Table is $cols by $rows "
# [ `echo $h1 | wc -w` = "2" ] || { echo "Error: $basename: Number of cols is not constant." >> $logdir/${basename}.log; echo "Number columns is not constant. Will exit";
if [ `echo $h1 | wc -w` != "2" ]; then
    mv pspp_out.csv tmp.pspp_out
    echo "" > pspp_out.csv
    while [ "`diff -q pspp_out.csv tmp.pspp_out`" ]; do

        echo `diff -q pspp_out.csv tmp.pspp_out` " "`wc -l pspp_out.csv`
        sed ':a;N;$!ba;s/

/g' tmp.pspp_out > pspp_out.csv
        sed ':a;N;$!ba;s/(\t")\n/1 /g' pspp_out.csv | sed ':a;N;$!ba;s/(\t"[^"]\t")+""\n/1 /g' | sed ':a;N;$!ba;s/(\t"[^"]\t")\n/1 /g' | sed ':a;N;$!ba;s/(\t"[^"]\t")\n/1 /g' | sed ':a;N;$!ba;s/(\t"[^"]\t")\n/1 /g'
        echo "Info: $basename: removed embedded newlines from dataset."
        echo "Info: $basename: removed embedded newlines from dataset." >> $logdir/${basename}.log;
    done
    h1=`awk -F"\t" '{print NF}' pspp_out.csv | uniq -c`
    if [ `echo $h1 | wc -w` != "2" ]; then
        echo " newlines remove leaves disarray "
        echo "Error: $basename: the newlines removal loop no helpee. Dataset in disarray." >> $logdir/${basename}.log;
        exit
    else
        echo -n " successfully removed newlines from pspp cells "
        echo "Info: $basename: removed newlines from table cells." >> $logdir/${basename}.log;
    fi
    tabse=`expr $cols - 2`
    wc -l st_out.csv
    echo strini
    sed ':a;N;$!ba;s/

/g' st_out.csv > tmp.st_out
    echo "" > st_out.csv
    while [ "`diff -q st_out.csv tmp.st_out`" ]; do
        mv tmp.st_out st_out.csv
        wc -l st_out.csv
        echo stround
        sed ':a;N;$!ba;s/(\n[^\\t\\n]*\\(\\t[^\\t\\n]*\\)\\{1,$tabse\\}\\}\\n/1 /" st_out.csv | sed ":a;N;$!ba;s/(\n[^\\t\\n]*\\(\\t[^\\t\\n]*\\)\\{1,$tabse\\}\\}\\n/1 /" | sed ":a;N;$!ba;s/(\n[^\\t\\n]*\\(\\t[^\\t\\n]*\\)\\{1,$tabse\\}\\}\\n/1 /"
    done
done
```

Format conversions

- Handle metadata

```
for i in `tail -n+2 ${fbase}-Vars_Lines | cut -d: -f1`; do
length=`expr $i - $startline - 1`
tail -n+$startline $indir/$file | head -n$length > cutfile
v_index=`grep -m1 "^ V_Index: " cutfile | cut -f2`
v_format=`grep -m1 "^ V_FormatS: " cutfile | cut -f2`
v_measure=`grep -m1 "^ V_Measure: " cutfile | cut -f2`
v_sql_datatype=`grep -m1 " V_Datatype: " cutfile | cut -f2 | sed -e 's/Integer/INT/' -e 's/String/VARCHAR/' -e 's/Decimal/DECIMAL/'`
v_sql_varlen=`grep -m1 "V_VarLen: " cutfile | cut -f2 | tr "." ","`
v_label=`grep -m1 "^ V_Label: " cutfile | cut -f2`
v_label_ddi=`echo "$v_label" | sed -e 's/\\&/\\&amp;/g' -e 's/</\\&lt;/g' -e 's/>/\\&gt;/g' -e 's/\\x27/\\&apos;/g' -e 's/"/\\&quot;/g'`
v_label_spss=`echo "$v_label" | sed "s/'/'/g"`
v_miss_spss=`grep -m1 "^ V_MissingR: " cutfile | cut -f2`
v_miss_d=`grep -m1 "^ V_MissingD: " cutfile | cut -f2`
v_sha=`grep -m1 "^ V_ValueL_sha: " cutfile | cut -f2`
if [[ $v_sha ]]; then
    sed "s/^(${v_sha})t.*)$/\\1${v_name} /" tmp.shalhelp > tmp.shalhelp
    mv tmp.shalhelp tmp.shalist
fi
echo -e " $v_name\\t$v_format" >> SPSS_var_form
echo -e " $v_name\\t$v_label_spss" >> SPSS_var_label
echo -e "label variable $v_name \\\"$v_label\\\">> Stata_var_label
echo -n " $v_name" >> tmp.M-$v_measure
[[ $v_miss_spss ]] && echo -e " $v_name\\t($v_miss_spss)" >> SPSS_missing
echo -e "\\t\\t<var ID=\\\"$v_name\\\" name=\\\"$v_name\\\">\\n\\t\\t\\t<labl>$v_label_ddi</labl>" >> DDI2_var
echo -e "\\v_index\\t$v_name\\t$v_label\\t$v_format\\t$v_measure\\t$v_miss_spss\\t$v_sha" >> tmp.sql-var_table
echo -e "\\t\\t$v_name $v_sql_datatype($v_sql_varlen)," >> tmp.sql-data_def
grep "^ " cutfile > tmp.vall_il
cut -f2 tmp.vall_il > tmp.vall_i
[ "`grep -m1 "^$v_sha " tmp.sql-vallab`" ] || sed "s/^/$v_sha/" tmp.vall_il >> tmp.sql-vallab
echo $v_miss_d | tr " " "\\n" > tmp.miss_d
for toto in `cat tmp.miss_d tmp.vall_i | sort -n | uniq`; do
    missus="N"
    if [ "`grep -x -m1 -- $toto tmp.miss_d`" ]; then
        missus="Y"
        echo -e "$v_name\\t$toto" >> tmp.sql-missings
    fi
    lab_lab=`grep -m1 "^ $toto " tmp.vall_il | cut -f3 | sed -e 's/\\&/\\&amp;/g' -e 's/</\\&lt;/g' -e 's/>/\\&gt;/g' -e 's/\\x27/\\&apos;/g' -e 's/"/\\&quot;/g'`
    echo -e "\\t\\t\\t<catgry missing=\\\"$missus\\\">\\n\\t\\t\\t\\t<trcatValu>$toto</catValu>\\n\\t\\t\\t\\t<labl>$lab_lab</labl>\\n\\t\\t\\t\\t<catgry>" >> DDI2_var
done
[ -s tmp.vall_il ] && mv tmp.vall_il tmp.vl-$v_sha
echo -e "\\t\\t</var>" >> DDI2_var
startline=$((i))
v_name=`grep "^$i:" ${fbase}-Vars_Lines | cut -d: -f3`
```

Provide access

- DBK download
 - ▶ in addition to the other formats
 - ▶ as a replacement for the archival version and downloads
- Enable direct access: API
 - ▶ machine-actionable
 - ▶ enables data linking
 - ▶ provides support for different formats via interface
 - ▶ crawling of data and metadata possible
 - ▶ makes visible what usage is needed
 - ▶ access regulations must be implemented

Build an easy to use API

- API should be RESTful
 - ▶ To enable good machine interaction

Example Action	Classic Webservice	RESTful
Get study list	GetStudies()	dbk/study/ HTTP Method GET
Get study description	GetDescription(studyID)	dbk/study/{studyID} HTTP Method GET
Get variables of study	GetVariables(studyID)	dbk/study/{studyID}/variable HTTP Method GET
Get variable description	GetVariable(studyID, var)	dbk/study/{study}/variable/{var} HTTP Method GET
Get variable data	GetData(studyID, var)	dbk/study/{study}/variable/{var}/data HTTP Method GET

Build an easy to use API

- API should be documented by OpenAPI
 - ▶ To enable easy programming against it

The OpenAPI Specification (fka The Swagger Specification)

build passing

Looking for the upcoming 3.0 version? See here:

[The pre-release OAS 3.0.0 Specification Branch](#)



The goal of The OpenAPI Specification is to define a standard, language-agnostic interface to REST APIs which allows both humans and computers to discover and understand the capabilities of the service without access to source code, documentation, or through network traffic inspection. When properly defined via OpenAPI, a consumer can understand and interact with the remote service with a minimal amount of implementation logic. Similar to what interfaces have done for lower-level programming, OpenAPI removes the guesswork in calling the service.

Use cases for machine-readable API interfaces include interactive documentation, code generation for documentation, client, and server, as well as automated test cases. OpenAPI-enabled APIs expose JSON files that correctly adhere to the OpenAPI Specification, documented in this repository. These files can either be produced and served statically, or be generated

Build an easy to use API

- API should be documented by OpenAPI
 - ▶ To enable easy programming against it

The screenshot shows the Swagger Petstore API documentation. At the top, there's a green header with the Swagger logo and a search bar containing the URL `http://petstore.swagger.io/v2/swagger.json`. Below the header, the title "Swagger Petstore" is displayed with a version badge "1.0.0". The main content area provides a description of the API, including a base URL and a sample API key. Below this, there are links for "Terms of service", "Contact the developer", "Apache 2.0", and "Find out more about Swagger". A "Schemes" dropdown menu is set to "HTTP", and an "Authorize" button is visible. The bottom section, titled "pet Everything about your Pets", lists five API endpoints: a POST endpoint for adding a new pet, a PUT endpoint for updating an existing pet, a GET endpoint for finding pets by status, a GET endpoint for finding pets by tags, and a GET endpoint for finding a pet by ID. Each endpoint is represented by a colored bar with its method, path, description, and a lock icon.

Build an easy to use API

 **swagger**

<http://localhost:1337/swagger/doc>

[Explore](#)

dbk-api

access social science research data of the GESIS Data Catalogue DBK

Created by zenk

Swagger

[Show/Hide](#) | [List Operations](#) | [Expand Operations](#)

Dataset

[Show/Hide](#) | [List Operations](#) | [Expand Operations](#)

Record

[Show/Hide](#) | [List Operations](#) | [Expand Operations](#)

Study

[Show/Hide](#) | [List Operations](#) | [Expand Operations](#)

[GET](#) /study/{id} [Read Object\(s\)](#)

[POST](#) /study/{id} [Create Object\(s\)](#)

[PUT](#) /study/{id} [Update Object\(s\)](#)

[GET](#) /study [Read Object\(s\)](#)

[POST](#) /study [Create Object\(s\)](#)

User

[Show/Hide](#) | [List Operations](#) | [Expand Operations](#)

[GET](#) /user/{id} [Read Object\(s\)](#)

[POST](#) /user/{id} [Create Object\(s\)](#)

[PUT](#) /user/{id} [Update Object\(s\)](#)

[POST](#) /login [Create Object\(s\)](#)

[POST](#) /signup [Create Object\(s\)](#)

Build an easy to use API

■ Prototype

- ▶ Using Sails.js, MVC web framework built on Node.js and EJS

dbk-api

Welcome

This is the API to access the social science research data from the GESIS Data Catalogue [DBK](#).

Learn more »

Account

1 You need an account with the DBK to use this API - [register](#). You may signup for a test here: [signup](#)

Login

2 Please [login](#) with your DBK or test account. Or [logout](#).

Use the dbk-api

3 You can now use the dbk-api to access your data.

View details »

© 2017 [GESIS](#), Wolfgang Zenk-Möltgen

WORK IN PROGRESS

<http://ejs.co/>

<https://nodejs.org/>

<http://sailsjs.com/>

Future use

- Manage dataset access with a DBK user account

Settings for DBK/datorium account

Account Settings for wolfgang.zenk-moeltgen@gesis.org

Account

First Name *

Last Name *

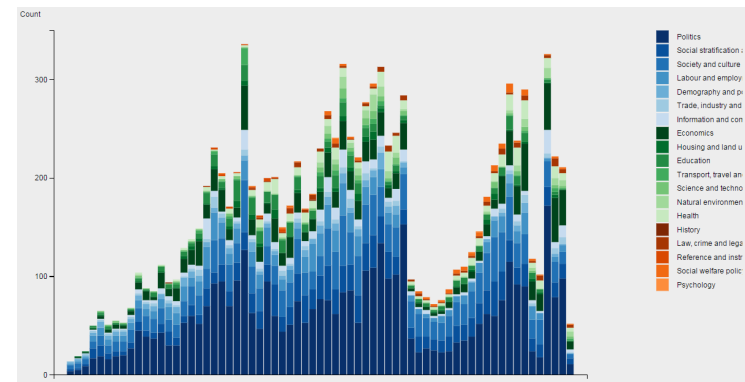
Country *

You can use the following studies via the DBK data API

ZA5760	Intermediate inquiry 2016 of Long-term Panel (GLES) http://dbk.gesis.org/api/v1/study/ZA5760/
ZA5300	Pre-election Cross Section (GLES 2009) http://dbk.gesis.org/api/v1/study/ZA5300/

Future use

- Use Case
 - ▶ Access the API to fetch data for analysis from any programming language
 - ▶ E.g. use the R packages httr or RCurl
 - ▶ Combine the data with data from other sources
- There are many tutorials available, e.g.
 - ▶ <https://www.r-bloggers.com/getting-data-from-an-online-source/>
 - ▶ <https://www.r-bloggers.com/accessing-apis-from-r-and-a-little-r-programming/>
- In the future, you can use the DBK API to analyze data!



Thank you for your attention

gesis

Leibniz Institute
for the Social Sciences

Member of the
Leibniz
Leibniz Association