Schöps, Marcus et al.

Article — Published Version

The WBdigital Database: A Digital Repository for the Historical DIW-Wochenbericht, 1928–1968

Jahrbücher für Nationalökonomie und Statistik - Journal of Economics and Statistics

This Version is available at:
http://hdl.handle.net/11108/633

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der
Leibniz-Gemeinschaft

Data Observer

Marcus Schöps*, Enrico Wedekind, Tobias Gebel, Andreas O. Kempf,
Peter Löwe, Luca Kohlhepp, Alexander Gehrke and Frank Puppe

The WBdigital Database: A Digital Repository for the Historical DIW-Wochenbericht, 1928–1968

Abstract: Economic and social scientists are increasingly interested in historical data, but many relevant sources are still available in analog form, limiting accessibility and research potential. This article introduces the WBdigital database, which aims to improve this situation. The database provides digital access to the DIW Wochenbericht (1928–1968), including its economic texts and time series data covering both German and international perspectives. These research data offer an ideal framework for investigating important research questions related to historical contexts and current economic challenges. Thus, we hope that the WBdigital database will be a valuable tool for researchers interested in economic history.

Keywords: economic history; interwar period; postwar period; database; Germany

JEL Classification: N01; N14

*Corresponding author: Marcus Schöps, Research Infrastructure Department, German Institute for Economic Research (DIW Berlin), Mohrenstraße 58, 10117 Berlin, Germany, E-mail: mschoeps@diw.de
Enrico Wedekind and Tobias Gebel, Research Infrastructure Department, German Institute for Economic Research (DIW Berlin), Mohrenstraße 58, 10117 Berlin, Germany, E-mail: ewedekind@diw.de (E. Wedekind), tgebel@diw.de (T. Gebel)
Andreas O. Kempf, Academic Services Department, ZBW – Leibniz Information Center for Economics, Düsternbrooker Weg 120, 24105 Kiel, Germany, E-mail: a.kempf@zbw.eu
Peter Löwe, IT & eScience, WZB Berlin Social Science Centre, Reichpietschufer 50, 10785 Berlin, Germany, E-mail: peter.loewe@wzb.eu
Luca Kohlhepp, Alexander Gehrke and Frank Puppe, Chair for Artificial Intelligence and Knowledge Systems, Institute of Computer Science, Wuerzburg University, Am Hubland, D-97074 Wuerzburg, Germany, E-mail: luca.kohlhepp@uni-wuerzburg.de (L. Kohlhepp), alexander.gehrke@uni-wuerzburg.de (A. Gehrke), frank.puppe@uni-wuerzburg.de (F. Puppe)

1 Introduction

Economists and social scientists are increasingly turning to historical research to analyze long-term economic developments, evaluate economic theories, and gain a deeper understanding of the past.[1] While digitization and modern computing have expanded access to historical data and the tools to analyze it, many relevant historical data sources remain in analog form, limiting their scientific utility and leaving significant research potential untapped (Hesse 2017).

This article describes the development and content of the WBdigital database, a project aimed at expanding access to digitized data for historical research in economics and other social sciences. The WBdigital database is the result of the project "Wochenbericht digital in Wort und Zahl – Digitale Bereitstellung der DIW-Wochenberichtsinhalte 1928 bis 1968 (WBdigital)."[2] The objective was to provide digital access to the extensive data contained in the historical DIW Wochenbericht (WB). Up to 1968, the WB offers a comprehensive and detailed account of the German and international economy, covering key economic episodes of the interwar and post-war periods through ongoing analyses, reports, and economic data (Fremdling and Stäglin 2008).

The paper is structured as follows: First, we provide a brief description of the WB and project background. This is followed by an outline of our innovative approach to digitization. Next, we present an overview of the structure, content, and accessibility of the database. Finally, future research perspectives and possibilities for further development are discussed.

2 Data Source and Project Background

The WBdigital database is based exclusively on the historical issues of the WB, published by the German Institute for Economic Research (DIW Berlin) between 1928

---

1 Abramitzky (2015) presents empirical evidence demonstrating a substantial increase in both the number and proportion of economic history articles published in the top five economics journals (AER, QJE, JPE, Econometrica, and Restud) between 1990 and 2014. Several scholars have discussed a significant portion of this literature, including Nunn (2009), Abramitzky (2015), and Cantoni and Yuchtman (2021).

2 The project is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) as part of the Scientific Library Services and Information Systems (LIS) funding program (digitization and indexing). WBdigital is a joint project of the German Institute for Economic Research (DIW Berlin), the University of Würzburg, Chair of Artificial Intelligence and Knowledge Systems, the Leibniz Information Center for Economics (Leibniz-Informationszentrum Wirtschaft, ZBW), and the Social Science Research Center Berlin (Wissenschaftszentrum Berlin für Sozialforschung, WZB).

and 1968.[3] Since 1928, the WB has provided regular coverage of current economic issues and contributed the Institute's research findings to the public debate.[4] Prior to 1968, the WB typically consisted of two to three articles in which the Institute briefly presented its economic, market, country, or sector analyses. Moreover, the WB was frequently accompanied by thematically structured statistical overviews (the so-called Zahlenübersichten, ZU), which provided detailed and comprehensive information on the domestic and foreign state of the economy based on a multitude of indicators. These indicators encompass a wide range of economic and financial data, including production, construction, employment, monetary aggregates, prices, securities markets, foreign trade, transport, and public finances.[5]

The consistent and detailed coverage of economic developments, along with the quality of the economic data, makes the WB a valuable resource for historical economic and social research (Buchheim 2003, 2008; Dimsdale, Horsewood, and Van Riel 2004; Fremdling 2010; Ritschl 1992; Spoerer 1997).[6] However, its analog format has significantly hindered widespread use due to the extensive research and data collection efforts required. In response to numerous user requests, DIW Berlin's Research Infrastructure Department aimed to remove these access barriers and make the WB and its data permanently accessible and usable via a digital portal in line with the FAIR principles.[7] In addition to full-text digitization and indexing, the project focused on the ZU and their economic data. Using a novel technical approach, the printed data tables of the ZU were digitized in a largely automated process and processed into long, consistent time series within thematically structured datasets.

3 Digitization

This section outlines the process of converting the printed volumes of the WB into digitized and processable research data. We began by digitizing all volumes from

---

3 DIW Berlin was founded in 1925 as the Institute for Business Cycle Research (Institut für Konjunkturforschung, IfK) and was renamed the German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung, DIW) in 1941.
4 For a comprehensive account of the historical development of the WB, refer to Krengel (1986) and Fremdling and Stäglin (2008).
5 The indicators listed in the ZU are based on the Institute's own calculations, as well as on government sources such as official statistics, authorities, and ministries, and non-governmental sources such as trade associations, newspapers, and other research institutes.
6 Tooze (2001, 103–48; 2017) discusses the quality and longevity of the Institute's data from the 1920s and 1930s, concluding that the innovative macroeconomic approach taken by the IfK, in collaboration with the Reich's Statistical Office, in collecting economic information provides a valuable and reliable source for contemporary economic research.
7 FAIR principles stand for Findability, Accessibility, Interoperability, and Reusability, which are guidelines designed to improve the management and sharing of research data.

1. Beilage zum Wochenbericht des Deutschen Instituts für Wirtschaftsforschung																Monatliche Zahlen-übersicht März 1958	C
25. Jahrgang		Berlin, den 9. Mai 1958									Nr. 19						
Gegenstand	Ge-biet*)	Einheit†)		1957												1958	
				Jan.	Febr.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.	Jan.	Febr.	März
Anzahl der Werktage:				26	24	26	24	25	22,2	27	26,9	25	27	24,4	24	25,7	24	26
Industr. Auftragseingang ¹⁹) (kalendermonatlich) ¹) ²) . . .	BRD	1954 = 100	D	144	137	149	141	152	134	147	136	142	157	148	133	140	132	·
Grundstoffindustrie	»	"	"	137	134	147	136	147	135	160	141	143	155	142	130	134	124	·
Investitionsgüterindustrie . . .	»	"	"	155	149	163	153	160	140	152	155	150	160	157	155	166	161	·
Verbrauchsgüterindustrie . . .	»	"	"	138	123	131	131	145	124	121	103	131	155	143	106	112	103	·
Ind. Bruttoproduktionsw. ³)	BRD	Mill. DM	S	15559	15153	16529	16209	16805	15250	16981	16781	17265	18343	17665	17011	°)16500	15488	·

Figure 1: Excerpt from an original scan of the Zahlenübersicht C (issue no. 19 of the DIW Wochenbericht 1959). Only 8 of 90 lines are shown.

1928 to 1968, supplementing any low-quality volumes in the DIW library with loans from other libraries.[8] In total, 10,900 pages were scanned. We then used commercial OCR software (ABBYY FineReader) to extract data from the digital copies. While OCR recognition of texts was highly successful, the ZU presented a greater challenge due to its complex tabular structure. The hierarchical structure with multiple headers and columns was more complex than conventional tables, and the inclusion of footnotes and non-numeric cell entries, such as quotation marks, added further complications to recognition and transcription (see Figure 1).

To address these challenges, we combined machine learning models with a knowledge-based approach to automatically transcribe data from the ZU. First, we used the open-source OCR tool PaddleOCR to recognize the numbers and text entries in each table cell, reconstructing the rows and columns.[9] The large number of entries, small distances between cells, and occasional optical distortions made it difficult to assign coordinates accurately. To solve these issues, we employed the DBSCAN local clustering algorithm to map individual table cells to neighboring ones.[10] The lines separating columns were additionally detected using a U-Net deep learning architecture. Finally, heuristics were applied to validate the transcription for each row and column. This approach achieved an overall accuracy rate of 99.8 %, indicating a high precision in identifying rows and columns. However, the remaining 0.2 % error rate suggests that approximately 14 % of the ZU may contain at least one incorrectly recognized row or column, meaning that the structure of 86 % of the ZU was correctly identified.

---

8 Unfortunately, not all volumes of the WB up to 1943 have been preserved in their original form, as many were damaged or lost during the Second World War. After the war, these volumes were reconstructed with the help of the Institute's staff. Today, many of these reconstructed volumes can be found not only at DIW Berlin but also in the collections of external specialist libraries, such as the ZBW in Kiel.

9 https://pypi.org/project/paddleocr/.

10 https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html.

In a second step, we transcribed the table cells of each ZU using OCR software. For numerical entries, we replaced PaddleOCR's transcription with Guppy, a self-developed open-source OCR tool specifically designed to recognize numbers, which significantly reduced the error rate. To validate the transcriptions, we compared them with subsequent transcriptions of the same issue, leveraging the regular publication frequency (weekly, monthly, quarterly). Successive ZUs usually differ only in the outer columns, with the left-hand column being replaced by a new column on the right-hand side. This redundancy allowed us to calculate a majority vote for each cell's transcription. Identical numerical entries in two or three transcriptions indicate high confidence, while discrepancies indicate low confidence, which is highlighted by colors. Consequently, the co-editors only needed to manually check the colored table cells using our table editor, KIETA (Kempf, Krug, and Puppe 2023), which also displays an image section of the original scan next to the OCR transcription to facilitate manual review (see Figure 2).

The post-correction of a ZU with approximately 900 automatically transcribed numerical entries took only 5–10 min with KIETA, achieving an error rate of 1.3 %, mainly due to footnote characters and incorrectly cut table cells. Excluding these errors, the OCR error rate drops to 0.1 %. For the transcription of textual table cells within a ZU, we used the PaddleOCR model with automatic post-correction based on a dictionary of typical text snippets. KIETA also offers export to various formats such as .csv resp. Excel (see Figure 3) or .html, which were used for further processing and import into the database.

To create long and consistent time series from the ZU, we aggregated the transcripts and their variables over time, following a source-oriented approach to create
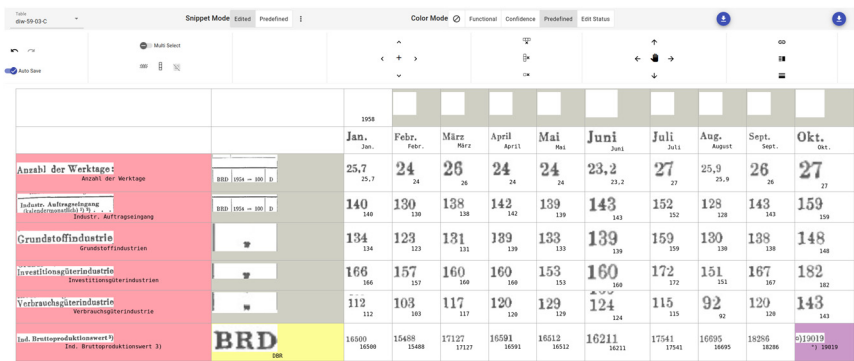

Figure 2: Editor KIETA for post-correction of the automatic transcription displays an image section for each cell, with the OCR-transcribed figures or text below it for easy correction by clicking. Cells marked in color indicate low confidence and require manual checking. Footnotes are marked in purple. The results can be saved and exported in various formats.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			1958											
2			Jan.	Febr.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
3	Anzahl der Werktage	BRD 1954 = 100 D	25,7	24	26	24	24	23,2	27	25,9	26	27	23,4	25
4	Industr. Auftragseingang		140	130	138	142	139	143	152	128	143	159	147	141
5	Grundstoffindustrien		134	123	131	139	133	139	159	130	138	148	133	134
6	Investitionsgüterindustrien		166	157	160	160	153	160	172	151	167	182	167	166
7	Verbrauchsgüterindustrie		112	103	117	120	129	124	115	92	120	143	140	115
8	Ind. Bruttoproduktionswert 3) BRD		16500	15488	17127	16591	16512	16211	17541	16695	18286	19019	17473	17627

Figure 3: Export from KIETA as a .csv file. The cell containing a footnote is highlighted in purple.

a digital surrogate of the original (Harvey and Press 1996).[11] This task was complicated by methodological and definitional changes over time, such as index revisions, territorial changes, and variations in measurement and currency units. Additionally, the ZU content was dynamic, with variables being added or removed over time. To address these challenges, we manually checked all ZUs for transformations affecting time series consistency and documented them.[12] Using this documentation, we segmented the aggregates of each individual variables into long, consistent time series based on standardized definitions and methods.[13]

4 The WBdigital Database

To present and provide access to the research data digitized and obtained from the WB, we created the WBdigital database. This database is freely accessible at https://wbdigital.diw.de, with a user interface available in both German and English.[14] The various sections of the database are described in more detail below.

4.1 *Wochenberichte* (Weekly Reports)

This section contains 35 volumes of the WB published between 1928 and 1968, including 1,542 issues and 3,451 articles, all of which can be viewed and downloaded as digital copies.[15] Full-text search is available using OCR recognition. Users can

---

11 In line with a source-oriented approach, we did not use interpolation or other statistical techniques to replace missing time series values, nor did we employ linking techniques like standardization or chain-linking to create long time series.
12 All footnotes and other information in the original data were also documented.
13 Due to overlapping time periods, users can link segments of a variable according to their own research interests.
14 Registration is required to download data from the database.
15 The publication of the WB was suspended in the winter of 1943 (16th volume) due to the war and resumed in early 1950 (17th volume).

search WB and their articles by title, year of publication, or topic (Figure 4).[16] For topic searches, keywords from the WB volumes have been mapped to the bilingual STW Thesaurus for Economics, extending search options with numerous German and English entries from the STW and concordances to other controlled vocabularies, such as the German Integrated Authority File (GND).[17]

Initially, a string match was performed using a Python script to compare entries in the keyword directories with preferred terms and synonym references in the STW. Based on this string comparison, intellectual mapping at the semantic level was carried out. Most entries in the keyword directories were mapped to one or more STW subject headings and linked to their persistent IDs. In several cases, alternative search entries or entirely new concept records were added to the thesaurus. For geographic locations below the state level, as well as for corporations and treaties, keyword entries were linked to corresponding GND entries to provide additional alternative search options. Time entries and brackets, such as those with formal descriptions, could not be mapped to the STW but can be searched using the full-text search.

4.2 *Zahlenübersichten* (Statistical Overviews)

This section contains transcripts of the various ZU series (A, B, C, D, E, F, Berlin, and Mitteldeutschland) produced during the digitization process (Figure 5). These
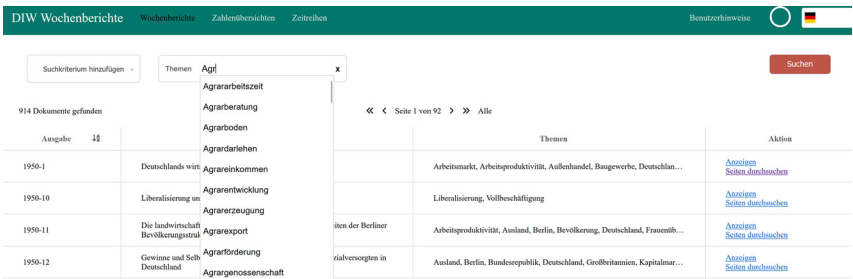
Figure 4: The topic search includes both the entries from the keyword indexes listed in the *Thema (WB + STW)* column and the subject headings from the STW, including concordances to other vocabularies. *DIW Wochenbericht* articles with the German keyword entry *"Landwirtschaft"* can be found with the English search entry *"agriculture",* among others.

---

16 Up to 1968, the articles in the DIW Wochenbericht were generally published without the author's name, except for the contributions by the first President Ernst Wagemann, which were marked with his name or the abbreviation E.W. It is therefore not possible to search for authors.

17 The acronym refers to the German name "Standard-Thesaurus Wirtschaft" see: http://zbw.eu/stw/version/latest/about.en.html.

DIW Wochenberichte   Wochenberichte   Zahlenübersichten   Zeitreihen   Benutzerhinweise   Original anzeigen

Jahr 1958   Reihe C

In Suchergebnis blättern : 3 von 5
Filter: J=ne=1958-1958   Zurück - Weiter   Zurück zur Suche

Blau hinterlegte Zellen enthalten nicht digitalisierte Fußnoten, siehe Original

Gegenstand	Gebiet *)	Einheit †)		Jan.	Febr.	März	April	Mai	Juni	Juli	August	Sept.	Okt.	Nov.	Dez.
Anzahl der Werktage:				25,7	24	26	24	24	23,2	27	25,9	26	27	23,4	25
Indust. Auftragseingang															
(kalendermonatlich) 1)2)	BRD	1954 = 100	D	140	130	138	142	139	143	152	128	143	159	147	141
Grundstoffindustrien	BRD	1954 = 100	D	134	123	131	139	133	139	159	130	138	148	133	134
Investitionsgüterindustrien	BRD	1954 = 100	D	166	157	160	160	153	160	172	151	167	182	167	166
Verbrauchsgüterindustrie	BRD	1954 = 100	D	112	103	117	120	129	124	115	92	120	143	140	115
Ind. Bruttoproduktionswert 3)	BRD	Mill. DM	S	16500	15488	17127	16591	16512	16211	17541	16695	18286	19019	17473	17627
Produktionsindex 4)															
Industrie insgesamt	BRD	1936 = 100		219	220	223	234	232	236	224	219	238	248	258	238
Bergbau	BRD	1936 = 100		157	153	153	155	155	155	148	145	149	152	158	146
Energieversorgungsbetriebe	BRD	1936 = 100		415	390	384	372	346	343	334	339	368	380	414	411
Bauhauptgewerbe	BRD	1936 = 100		112	117	128	210	236	248	234	232	237	233	243	202
Grundstoffe u. Produktionsgüter	BRD	1936 = 100		196	199	200	212	213	218	211	207	215	217	224	203
Investitionsgüter	BRD	1936 = 100		285	295	300	311	306	321	287	277	315	320	337	319
Verbrauchsgüter	BRD	1936 = 100		210	206	209	216	203	199	195	190	218	227	233	204
Nahrungs- und Genußmittel	BRD	1936 = 100		195	190	192	210	216	219	212	211	217	252	263	248
Industrieprodukt. West-Berlins 5)	W-B	1936 = 100		113	121	121	125	118	118	108	118	133	133	138	131
Produktion															
Eisenerz	BRD	1000 t	S	1640	1522	1607	1531	1527	1480	1402	1399	1431	1536	1482	1424

Jahr 1958
Reihe C
Titel Wochenbericht Zahlenübersicht C 1958
Short-Link https://diw-berichte.infor...

Anhänge:
- CSV /resources/pdf/tables /diw-59-03-C.csv

In Suchergebnis blättern : 3 von 5
Filter: J=ne=1958-1958   Zurück - Weiter   Zurück zur Suche

Figure 5: Presentation of a scrollable transcript of a ZU, based on Figure 1 in the web interface of the DIW-viewer. The four header columns, typically consistent for tables of the same type, were added. Figures with blue background color contain footnotes. Their meaning can be checked by clicking the button "original anzeigen" (show original). Download is available in CSV format.

transcripts provide an initial annual perspective on the ZU data and can be searched by variable, year, ZU series, and full text. The transcripts can be viewed in detail and downloaded in CSV format.

Please note that the transcripts display only the variables and their numerical values, not the footnotes from the original documents. However, footnotes are marked in the transcripts (by numbers or color highlighting) and can be checked through the linked original documents. Researchers merging data from different transcripts of the same series should be aware that definitional and methodological changes between years may affect the consistency of the data. In general, it is highly recommended to refer to the original sources when working with data from the transcripts.

4.3 *Zeitreihen* (Time Series)

The time series section of our database includes all digitized and processed time series from the ZU, which regularly accompanied the WB. It features around 2,500 mostly disaggregated, consistent macroeconomic and macro-financial series, covering various periods from 1926 to 1968. These series are organized into 16

Table 1: Thematic and temporal coverage of the time series datasets.

Dataset	Topics	Interwar coverage	Postwar coverage
Wochenzahlen	General economic data (weekly frequency)	X	X
Monatszahlen	General economic data (monthly frequency)	X	
A	Prices, interest rates, securities markets, central banking, private banking, credits, government finances, wages	X	X
B	Foreign trade, domestic trade, transportation	X	X
C	Production, consumption	X	X
D	Labor market, employment	X	X
E	General economic data for the USA, Great Britain, France, Italy, Poland (up to 1939), Netherlands, Switzerland	X	X
F	General economic data for the USSR, Bulgaria, Poland, Czechoslovakia, Romania, Hungary, Yugoslavia		X
Berlin	General economic data for West-Berlin		X
Mitteldeutschland	General economic data for the German Democratic Republic (including East-Berlin)		X

datasets that reflect the thematic and historical structure of the ZU (see Table 1) and can be accessed and downloaded via the database.

Each dataset is supported by an index table that facilitates quick and easy data searches. The index table provides information about the variables listed in a dataset, including time series frequency, time span, unit of measurement, author's notes, and references to the ZU used to create the time series. Additionally, each time series in an index table is assigned a code linked to the actual data section, simplifying navigation between information and data sections within a dataset.[18]

Due to the suspension of the WB during the war and the immediate post-war years, no data is available for this period. Consequently, most datasets are divided into interwar and postwar periods. The datasets primarily cover Germany, including comprehensive coverage for West Berlin and the GDR, but also include some data for other countries (see Table 1).

Most time series are available at monthly and, to a lesser extent, quarterly frequencies. Additionally, based on the ZU series of the "Wochenzahlen" (weekly figures) published from 1928 to 1961, we digitized a significant number of high-frequency

---

18 The concept and functionality of the index tables are inspired by a similar index table developed by Costache, Dimitrova, and Lazaretou (2015).

weekly time series for various economic areas, primarily covering Germany, but also including data for Great Britain and the United States. Weekly time series are rare in historical economic research today, yet they allow for more detailed and sophisticated analysis that less frequent data may obscure. Thus, the higher granularity of weekly time series can provide new insights into short-term reactions of economic and financial markets, such as responses to crises, policy changes, sectoral dynamics, consumer behavior shifts, and financial market volatility.

To ensure long-term usability for various research interests, each dataset is accompanied by data documentation. This documentation, also accessible through the database, provides descriptive information on the data and the preparation process, as well as historical context for many variables, including explanations of survey and calculation methods (and their revisions), historical terminology, units of measurement, and currencies.

5 Conclusions

The historical WB (1928–1968) is an impressive collection of economic information. The digital processing of these records has significantly facilitated access to this important source for historical research in economics and social sciences, while also expanding the availability of historical economic data. The digitized data encompass pivotal periods in German and international economic history, providing a framework for investigating significant research questions relevant to understanding both historical contexts and current economic, political, and social challenges.

Historical economic texts on German and international economic developments are now available in full text, searchable using various criteria in both German and English. OCR recognition not only enables full-text search but also allows further processing using digital methods such as text mining, providing new impetus for the application of these approaches in economics and economic history.

Additionally, long and consistent time series were digitized from the regularly attached ZU, significantly enhancing the availability of macroeconomic and macro-financial data for the interwar and postwar periods. A source-oriented approach ensures these data are broadly applicable, supporting comparative analyses, long-term trend investigations, studies on political and economic interactions, international trade relations, and the effects of economic policies. In particular, the provision of weekly time series opens new possibilities for long-term macroeconomic and economic historical research.

From a technical perspective, an innovative approach has been developed for the time- and cost-efficient digitization and transcription of large volumes of printed

data from complex table formats. This approach is expected to benefit future digitization projects, making more relevant historical data sources digitally accessible.[19]

WBdigital represents an important first step by DIW Berlin's Research Infrastructure Department in providing digital access to historical research data. In the future, further relevant data sources from DIW Berlin's historical archives, such as the Vierteljahrshefte zur Konjunkturforschung, may also be digitized, thereby creating a comprehensive and permanent digital infrastructure for historical economic data at DIW Berlin.

References

Abramitzky, Ran. 2015. *Economics and the Modern Economic Historian*. 21636. NBER Working Papers. Cambridge: National Bureau of Economic Research.

Buchheim, Christoph. 2003. "Die Erholung von Der Weltwirtschaftskrise 1932/33 in Deutschland." *Jahrbuch Für Wirtschaftsgeschichte/Economic History Yearbook* 44 (1). https://doi.org/10.1524/jbwg.2003.44.1.13.

Buchheim, Christoph. 2008. "Das NS-Regime Und Die Überwindung Der Weltwirtschaftskrise in Deutschland." *Vierteljahrshefte Für Zeitgeschichte* 56 (3): 381–414.

Cantoni, Davide, and Noam Yuchtman. 2021. "Historical Natural Experiments: Bridging Economics and Economic History." In *The Handbook of Historical Economics*, 213–41. London, San Diego, Cambridge, Oxford: Elsevier.

Costache, Brindusa, Kalina Dimitrova, and Sophia M. Lazaretou. 2015. "South-Eastern European Monetary and Economic Statistics from the Nineteenth Century to WWII." *eabh Papers* 15-02. The European Association for Banking and Financial History (EABH).

Dimsdale, Nicholas, Nicholas Horsewood, and Arthur Van Riel. 2004. "Real Wages and Unemployment in Weimar Germany." *Oxford Economic and Social History Working Papers* 056. Oxford: University of Oxford, Department of Economics.

Fremdling, Rainer. 2010. "German Agricultural Employment, Production and Labour Productivity: A New Benchmark for 1936 and a Note on Hoffmann's Tales." *Jahrbuch Für Wirtschaftsgeschichte/Economic History Yearbook* 51 (1): 215–28.

Fremdling, Rainer, and Reiner Stäglin. 2008. "Profund, Präzise, Pünktlich: 80 Jahre Wochenbericht Spiegeln Die Deutsche Wirtschaftsgeschichte." *DIW-Wochenbericht* 75 (14): 168–74.

Harvey, Charles, and Jon Press. 1996. *Databases in Historical Research*. London: Macmillan Education UK.

Hesse, Jan-Otmar. 2017. "Verborgene Schätze der Wirtschaftsgeschichte." *Fazit – Das Wirtschaftsblog (blog)*. https://blogs.faz.net/fazit/2017/11/23/nicht-geborgene-schaetze-der-wirtschaftsgeschichte-9379/#comments (accessed November 23, 2017).

---

19 The methods developed for digitizing and transcribing the ZU will be published as open-source software for subsequent use.

Kempf, Sebastian, Markus Krug, and Frank Puppe. 2023. "KIETA: Key-Insight Extraction from Scientific Tables." *Applied Intelligence* 53 (8): 9513–30.

Krengel, Rolf. 1986. *Das Deutsche Institut Für Wirtschaftsforschung (Institut Für Konjunkturforschung)*, 1925–79. Berlin: Duncker & Humblot.

Nunn, Nathan. 2009. "The Importance of History for Economic Development." *Annual Review of Economics* 1 (1): 65–92.

Ritschl, Albrecht. 1992. "Über Die Höhe Und Struktur der Gesamtwirtschaftlichen Investitionen in Deutschland 1935–38." *Vierteljahrschrift Für Sozial- Und Wirtschaftsgeschichte* 79 (2): 156–76.

Spoerer, M. 1997. "Weimar's Investment and Growth Record in Intertemporal and International Perspective." *European Review of Economic History* 1 (3): 271–97.

Tooze, J. Adam. 2001. *Statistics and the German State, 1900–1945: The Making of Modern Economic Knowledge. Cambridge Studies in Modern Economic History 9*. Cambridge: Cambridge University Press.

Tooze, J. Adam. 2017. "Wirtschaftsstatistik Im Reichswirtschaftsministerium, in Seinem Statistischen Reichsamt Und Im Institut Für Konjunkturforschung." In *Wirtschaftspolitik in Deutschland 1917–1990*. 1st ed. 361–420. Boston: Walter de Gruyter.