

# Company ID Linktables (IDLINK)

## Data Report 2023-15

Data available from 1987-01-01 to 2023-01-31

Version: 2023-1-6

DOI: 10.12757/Bbk.IDLINK.2023.1.6

Deutsche Bundesbank, Research Data and Service Centre

Eniko Gábor-Tóth

Christopher-Johannes Schild

Susanne Walter

## Abstract

This data report describes the research dataset “Company ID Linktables”, the company ID-linkage tables produced by the RDSC company data record linkage<sup>1)</sup>, using a structured metadata schema.<sup>2)</sup>

**Keywords:** Company data, ID Linkage Tables, Record Linkage, Data Matching

**Version:** 2023-1-6

**DOI:** 10.12757/Bbk.IDLINK.2023.1.6

**Citation:** Gábor-Tóth, E., Schild, C., and Walter, S. (2023). Company ID Linktables (IDLINK), Data Report 2023-15 – Version 2023-1-6. Deutsche Bundesbank, Research Data and Service Centre.

---

<sup>1</sup> The department responsible for RIAD provided additional ID-mappings between RIAD and USTAN, RIAD and JANIS, and between RIAD and BAKIS-M. These tables also entered the ID-linkage tables in IDLINK: they were consolidated with the matching result from the RDSC record linkage in order to further improve the matching rate.

<sup>2</sup> We thank Miro Holzer and Alice Micheler for providing additional match information produced by the RIAD Team of Deutsche Bundesbank. The metadata scheme is derived from the “Data Documentation Initiative” (DDI, <http://www.ddialliance.org>).

## Contents

<b>1 Dataset description</b>	<b>4</b>
1.1 Overview and identification	4
1.2 Dataset scope and coverage	4
1.3 Data appraisal	5
1.4 Data accessibility	6
1.5 Files Description	7
<b>2 Description of Variables</b>	<b>10</b>
2.1 Overview of Variables	10
2.2 Details of Variables	10
<b>References</b>	<b>15</b>

# 1 Dataset description

## 1.1 Overview and identification

The “Company ID Linktables” dataset comprises 79 ID linkage tables. Each of these 79 tables allows for linking two research (analytical) datasets. Through the pairwise linkage of BVD, MiDi, URS, USTAN, JANIS, BAKIS-M, LEI, RIAD (AnaCredit), SIFCT, SITS, and Schufa, “Company ID Linktables” allows for jointly evaluating these datasets for various time periods.

The individual ID linkage tables are two-column tables containing a different company identifier in each column, resulting in a table of ID value pairs, showing which ID values of the two IDs refer to the same real-world company entity. That is, the ID linkage tables contain the list of native ID pairs (or anonymized versions thereof) that form a match between two datasets. The need for these ID linkage tables originates from the fact that company data is often held in separate databases, which use different company identifiers. The ID linkage tables allow linking different company datasets.

The tables are produced by the RDSC through the use of current record linkage techniques, which, among other methods, include comprehensive data cleaning, matching based on common external IDs, as well as name and place based matching, which includes probabilistic matching using supervised machine learning. The technical properties of this record linkage process are described by Gabor-Toth, Schild, and Walter (2023a).

The department responsible for RIAD provided additional ID-mappings between RIAD and USTAN, RIAD and JANIS, and between RIAD and BAKIS-M. These tables also entered the ID-linkage tables in IDLINK: they were consolidated with the matching result from the RDSC record linkage in order to further improve the matching rate.

The ID value pairs included in these linkage table define overlaps between the different company datasets. These overlaps are described by Gabor-Toth, Schild, and Walter (2023b).

## 1.2 Dataset scope and coverage

### Legal framework

The Research Data and Service Centre (RDSC) is mandated by the Deutsche Bundesbank to use internal company master data for linkage purposes, to provide linked company data to internal analysts and anonymized linked company data to internal and external researchers.

### Unit of analysis

The unit of analysis is a single company. Generally this refers to legal entities. In each dataset to be linked, entities are represented by their dataset-specific native identifier.

## Time periods

The time periods covered by the data depend on the time periods of the input data sources. The earliest time references for ID values in the input data stem from 1979, the latest stem from 2023. For details see Gabor-Toth et al. (2023b)

## Geographic coverage

Germany

## Universe

The universes of the input datasets of our record linkage define the universe of our ID linkage tables. This universe covers all companies that appear at least once in any of the master datasets that have entered the record linkage. For details on the datasets' universes, we refer to the respective datasets' official documentation. For an overview on these datasets' universes and for their respective documentations see Gabor-Toth et al. (2023b)

## Historical changes

The input datasets' universes change over time, for example due to changes of reporting thresholds, or quality of data historization. Data integration for historic company data at Deutsche Bundesbank is in its infancy, understanding the different company databases and their universes and how these universes change over time is one of the challenges. For an attempt to summarize the current state of knowledge on these matters see Gabor-Toth et al. (2023b)

## 1.3 Data appraisal

### Quality checks

Quality checks are implemented at several points during the record linkage process. For example, before entering the record linkage, certain quality checks are undertaken for the input data, such as examining the filling ratios for positions that are relevant for linking the datasets. The ID-linkage tables are analyzed for match prediction precision and recall / match coverage, see Gabor-Toth et al. (2023a). Match coverage is further analyzed by Gabor-Toth et al. (2023b), with a focus on discussing the plausibility of the implied sizes of the data universe overlaps.

### Data editing

Before entering the record linkage that produced the ID-linkage tables described in this report, the input data is standardized by the data specialist of the RDSC responsible for the dataset. Standardization

occurs according to the data standard defined by AnaCredit RIAD, currently v2.1.<sup>3)</sup>

## Data Anonymization

Before the ID linkage tables are provided to researchers, all external company identifiers are anonymized using a secure hash algorithm (SHA256) as recommended by the German Federal Institute for Information Security (Bundesamt für Sicherheit in der Informationstechnik, 2020).<sup>4)</sup>

## 1.4 Data accessibility

Data from external sources such as BvD can only be accessed if the researcher has a licence with the data provider that allows the use of their data. Some datasets that have no research dataset counterpart cannot be accessed by external users, such as the case for the URS dataset, which we are allowed to use only for statistical purposes such as data linkage, but not for generating research data. The Schufa data can only be accessed by internal researchers and analysts.

## Research proposal conditions

A research proposal is checked for feasibility of the research project given the research data, i.e. the suitability of the data to answer the research questions raised by the proposal. The research project must be of public interest, that is without commercial goals.

## Institutional access conditions

The researcher must be affiliated with a research institution that clearly has a scientific, noncommercial agenda.

## Contact

Deutsche Bundesbank, Research Data and Service Centre (RDSC)  
E-mail: [fdsz-data@bundesbank.de](mailto:fdsz-data@bundesbank.de)  
Homepage: <https://www.bundesbank.de/rdsc>

## Deposit requirements

The researcher must sign a confidentiality agreement and a contract with the Deutsche Bundesbank. To use microdata available in BAKIS-M a separate contract needs to be signed between the research institute that the researcher is affiliated with and the Deutsche Bundesbank

---

<sup>3</sup> <https://www.bundesbank.de/de/service/meldewesen/bankenstatistik/formate-xml/formate-xml--611846>

<sup>4</sup> To enable the researcher to use the ID-mappingtables to link different datasets, the IDs contained in the research data are hashed using the same hash algorithm.

## Citation requirements

For any study or other document which is made available to the public and contains information derived from the provided data, the researcher is obliged to properly cite the data source as:

Gábor-Tóth, E., Schild, C., and Walter, S. (2023). Company ID Linktables (IDLINK), Data Report 2023-15 – Version 2023-1-6. Deutsche Bundesbank, Research Data and Service Centre.

## 1.5 Files Description

The files contain ID value pairs, showing which ID values of the two IDs refer to the same real-world company entity. The 54 files in Table 1, 'Files - ID LINK 'not anonymized'', on page 8, contain non-anonymized identifiers and are only available to internal analysts. The 25 files in Table 2, 'Files - ID LINK 'anonymized'', on page 9, only contain anonymized identifiers and are (in principle, depending on data access conditions for the respective datasets) also available to data users that may only use anonymized data.

### File Structure

The files contain 2 columns (two IDs).

### Type of Files

ASCII

### Data Formats

comma-delimited

Table 1: Files - ID LINK 'not anonymized'

Filename	IDs	Datasets
*orig_v2023-1-7_BBK_HM_CO_ID2_AWMUS_CD*	AWMUS_CD - BBK_HM_CO_ID2	AWMuS - JANIKA
*orig_v2023-1-7_BBK_HM_CO_ID2_BVD_CD*	BVD_CD - BBK_HM_CO_ID2	BVD - JANIKA
*orig_v2023-1-7_BBK_HM_CO_ID2_DE_BAKISN_CD*	DE_BAKISN_CD - BBK_HM_CO_ID2	BAKIS-M - JANIKA
*orig_v2023-1-7_BBK_HM_CO_ID2_DE_DESTATIS_CD_STBL*	DE_DESTATIS_CD_STBL - BBK_HM_CO_ID2	URS - JANIKA
*orig_v2023-1-7_BBK_HM_CO_ID2_ENTTY_RIAD_CD*	ENTTY_RIAD_CD - BBK_HM_CO_ID2	AnaCredit - JANIKA
*orig_v2023-1-7_BBK_HM_CO_ID2_ISIN*	ISIN - BBK_HM_CO_ID2	ISIN - JANIKA
*orig_v2023-1-7_BVD_CD_AWMUS_CD*	BVD_CD - AWMUS_CD	BVD - AWMuS
*orig_v2023-1-7_DE_BAKISN_CD_AWMUS_CD*	DE_BAKISN_CD - AWMUS_CD	BAKIS-M - AWMuS
*orig_v2023-1-7_DE_BAKISN_CD_BVD_CD*	DE_BAKISN_CD - BVD_CD	BAKIS-M - BVD
*orig_v2023-1-7_DE_DESTATIS_CD_STBL_AWMUS_CD*	AWMUS_CD - DE_DESTATIS_CD_STBL	AWMuS - URS
*orig_v2023-1-7_DE_DESTATIS_CD_STBL_BVD_CD*	DE_DESTATIS_CD_STBL - BVD_CD	URS - BVD
*orig_v2023-1-7_DE_DESTATIS_CD_STBL_DE_BAKISN_CD*	DE_DESTATIS_CD_STBL - DE_BAKISN_CD	URS - BAKIS-M
*orig_v2023-1-7_ENTTY_RIAD_CD_AWMUS_CD*	ENTTY_RIAD_CD - AWMUS_CD	AnaCredit - AWMuS
*orig_v2023-1-7_ENTTY_RIAD_CD_BVD_CD*	ENTTY_RIAD_CD - BVD_CD	AnaCredit - BVD
*orig_v2023-1-7_ENTTY_RIAD_CD_DE_BAKISN_CD*	ENTTY_RIAD_CD - DE_BAKISN_CD	AnaCredit - BAKIS-M
*orig_v2023-1-7_ENTTY_RIAD_CD_DE_DESTATIS_CD_STBL*	ENTTY_RIAD_CD - DE_DESTATIS_CD_STBL	AnaCredit - URS
*orig_v2023-1-7_ISIN_AWMUS_CD*	ISIN - AWMUS_CD	ISIN - AWMuS
*orig_v2023-1-7_ISIN_BVD_CD*	ISIN - BVD_CD	ISIN - BVD
*orig_v2023-1-7_ISIN_DE_BAKISN_CD*	ISIN - DE_BAKISN_CD	ISIN - BAKIS-M
*orig_v2023-1-7_ISIN_DE_DESTATIS_CD_STBL*	ISIN - DE_DESTATIS_CD_STBL	ISIN - URS
*orig_v2023-1-7_ISIN_ENTTY_RIAD_CD*	ISIN - ENTTY_RIAD_CD	ISIN - AnaCredit
*orig_v2023-1-7_JANIS_CD_AWMUS_CD*	JANIS_CD - AWMUS_CD	JANIS - AWMuS
*orig_v2023-1-7_JANIS_CD_BVD_CD*	JANIS_CD - BVD_CD	JANIS - BVD
*orig_v2023-1-7_JANIS_CD_DE_BAKISN_CD*	JANIS_CD - DE_BAKISN_CD	JANIS - BAKIS-M
*orig_v2023-1-7_JANIS_CD_DE_DESTATIS_CD_STBL*	JANIS_CD - DE_DESTATIS_CD_STBL	JANIS - URS
*orig_v2023-1-7_JANIS_CD_ENTTY_RIAD_CD*	JANIS_CD - ENTTY_RIAD_CD	JANIS - AnaCredit
*orig_v2023-1-7_JANIS_CD_ISIN*	JANIS_CD - ISIN	JANIS - ISIN
*orig_v2023-1-7_LEI_AWMUS_CD*	LEI - AWMUS_CD	LEI - AWMuS
*orig_v2023-1-7_LEI_BBK_HM_CO_ID2*	LEI - BBK_HM_CO_ID2	LEI - JANIKA
*orig_v2023-1-7_LEI_BVD_CD*	LEI - BVD_CD	LEI - BVD
*orig_v2023-1-7_LEI_DE_BAKISN_CD*	LEI - DE_BAKISN_CD	LEI - BAKIS-M
*orig_v2023-1-7_LEI_DE_DESTATIS_CD_STBL*	LEI - DE_DESTATIS_CD_STBL	LEI - URS
*orig_v2023-1-7_LEI_ENTTY_RIAD_CD*	LEI - ENTTY_RIAD_CD	LEI - AnaCredit
*orig_v2023-1-7_LEI_ISIN*	LEI - ISIN	LEI - ISIN
*orig_v2023-1-7_LEI_JANIS_CD*	LEI - JANIS_CD	LEI - JANIS
*orig_v2023-1-7_SCHUFA_ID_AWMUS_CD*	SCHUFA_ID - AWMUS_CD	SCHUFA - AWMuS
*orig_v2023-1-7_SCHUFA_ID_BBK_HM_CO_ID2*	SCHUFA_ID - BBK_HM_CO_ID2	SCHUFA - JANIKA
*orig_v2023-1-7_SCHUFA_ID_BVD_CD*	SCHUFA_ID - BVD_CD	SCHUFA - BVD
*orig_v2023-1-7_SCHUFA_ID_DE_BAKISN_CD*	SCHUFA_ID - DE_BAKISN_CD	SCHUFA - BAKIS-M
*orig_v2023-1-7_SCHUFA_ID_DE_DESTATIS_CD_STBL*	SCHUFA_ID - DE_DESTATIS_CD_STBL	SCHUFA - URS
*orig_v2023-1-7_SCHUFA_ID_ENTTY_RIAD_CD*	SCHUFA_ID - ENTTY_RIAD_CD	SCHUFA - AnaCredit
*orig_v2023-1-7_SCHUFA_ID_ISIN*	SCHUFA_ID - ISIN	SCHUFA - ISIN
*orig_v2023-1-7_SCHUFA_ID_JANIS_CD*	SCHUFA_ID - JANIS_CD	SCHUFA - JANIS
*orig_v2023-1-7_SCHUFA_ID_LEI*	SCHUFA_ID - LEI	SCHUFA - LEI
*orig_v2023-1-7_USTAN_CD_AWMUS_CD*	USTAN_CD - AWMUS_CD	USTAN - AWMuS
*orig_v2023-1-7_USTAN_CD_BBK_HM_CO_ID2*	USTAN_CD - BBK_HM_CO_ID2	USTAN - JANIKA
*orig_v2023-1-7_USTAN_CD_BVD_CD*	USTAN_CD - BVD_CD	USTAN - BVD
*orig_v2023-1-7_USTAN_CD_DE_BAKISN_CD*	USTAN_CD - DE_BAKISN_CD	USTAN - BAKIS-M
*orig_v2023-1-7_USTAN_CD_DE_DESTATIS_CD_STBL*	USTAN_CD - DE_DESTATIS_CD_STBL	USTAN - URS
*orig_v2023-1-7_USTAN_CD_ENTTY_RIAD_CD*	USTAN_CD - ENTTY_RIAD_CD	USTAN - AnaCredit
*orig_v2023-1-7_USTAN_CD_ISIN*	USTAN_CD - ISIN	USTAN - ISIN
*orig_v2023-1-7_USTAN_CD_JANIS_CD*	USTAN_CD - JANIS_CD	USTAN - JANIS
*orig_v2023-1-7_USTAN_CD_LEI*	USTAN_CD - LEI	USTAN - LEI
*orig_v2023-1-7_USTAN_CD_SCHUFA_ID*	USTAN_CD - SCHUFA_ID	USTAN - SCHUFA

<sup>a</sup> AWMuS holds the master data for MiDi, SITs and SIFCT. Therefore, through AWMuS, datasets can be linked to the three Bundesbank research datasets related to international trade statistics.



Table 2: Files - ID LINK 'anonymized'

Filename	IDs	Datasets
*anon_v2023-1-7_BVD_CD_AWMUS_CD*	anonym_BVD_CD - AWMUS_CD	BVD - AWMuS
*anon_v2023-1-7_ENTTY_RIAD_CD_AWMUS_CD*	anonym_ENTTY_RIAD_CD - AWMUS_CD	AnaCredit - AWMuS
*anon_v2023-1-7_ENTTY_RIAD_CD_BVD_CD*	anonym_ENTTY_RIAD_CD - anonym_BVD_CD	AnaCredit - BVD
*anon_v2023-1-7_ISIN_AWMUS_CD*	anonym_ISIN - AWMUS_CD	ISIN - AWMuS
*anon_v2023-1-7_ISIN_ENTTY_RIAD_CD*	anonym_ISIN - anonym_ENTTY_RIAD_CD	ISIN - AnaCredit
*anon_v2023-1-7_JANIS_CD_AWMUS_CD*	BBK_HM_CO_ID1 - AWMUS_CD	JANIS - AWMuS
*anon_v2023-1-7_JANIS_CD_BVD_CD*	BBK_HM_CO_ID1 - anonym_BVD_CD	JANIS - BVD
*anon_v2023-1-7_JANIS_CD_ENTTY_RIAD_CD*	BBK_HM_CO_ID1 - anonym_ENTTY_RIAD_CD	JANIS - AnaCredit
*anon_v2023-1-7_JANIS_CD_ISIN*	BBK_HM_CO_ID1 - anonym_ISIN	JANIS - ISIN
*anon_v2023-1-7_LEI_AWMUS_CD*	anonym_LEI - AWMUS_CD	LEI - AWMuS
*anon_v2023-1-7_LEI_ENTTY_RIAD_CD*	anonym_LEI - anonym_ENTTY_RIAD_CD	LEI - AnaCredit
*anon_v2023-1-7_LEI_JANIS_CD*	anonym_LEI - BBK_HM_CO_ID1	LEI - JANIS
*anon_v2023-1-7_SCHUFA_ID_AWMUS_CD*	anonym_SCHUFA_ID - AWMUS_CD	SCHUFA - AWMuS
*anon_v2023-1-7_SCHUFA_ID_BVD_CD*	anonym_SCHUFA_ID - anonym_BVD_CD	SCHUFA - BVD
*anon_v2023-1-7_SCHUFA_ID_ENTTY_RIAD_CD*	anonym_SCHUFA_ID - anonym_ENTTY_RIAD_CD	SCHUFA - AnaCredit
*anon_v2023-1-7_SCHUFA_ID_ISIN*	anonym_SCHUFA_ID - anonym_ISIN	SCHUFA - ISIN
*anon_v2023-1-7_SCHUFA_ID_JANIS_CD*	anonym_SCHUFA_ID - BBK_HM_CO_ID1	SCHUFA - JANIS
*anon_v2023-1-7_SCHUFA_ID_LEI*	anonym_SCHUFA_ID - anonym_LEI	SCHUFA - LEI
*anon_v2023-1-7_USTAN_CD_AWMUS_CD*	anonym_USTAN_CD - AWMUS_CD	USTAN - AWMuS
*anon_v2023-1-7_USTAN_CD_BVD_CD*	anonym_USTAN_CD - anonym_BVD_CD	USTAN - BVD
*anon_v2023-1-7_USTAN_CD_ENTTY_RIAD_CD*	anonym_USTAN_CD - anonym_ENTTY_RIAD_CD	USTAN - AnaCredit
*anon_v2023-1-7_USTAN_CD_ISIN*	anonym_USTAN_CD - anonym_ISIN	USTAN - ISIN
*anon_v2023-1-7_USTAN_CD_JANIS_CD*	anonym_USTAN_CD - BBK_HM_CO_ID1	USTAN - JANIS
*anon_v2023-1-7_USTAN_CD_LEI*	anonym_USTAN_CD - anonym_LEI	USTAN - LEI
*anon_v2023-1-7_USTAN_CD_SCHUFA_ID*	anonym_USTAN_CD - anonym_SCHUFA_ID	USTAN - SCHUFA

<sup>a</sup> AWMuS holds the master data for MiDi, SITS and SIFCT. Therefore, through AWMuS, datasets can be linked to the three Bundesbank research datasets related to international trade statistics.

## 2 Description of Variables

### 2.1 Overview of Variables

Name	Label
AWMUS_CD	AWMuS identifier. Anonymous. Original name: MLD_NR
BBK_HM_CO_ID1	JANIS identifier (anonymized)
BBK_HM_CO_ID2	JANIKA identifier (anonymized)
BVD_CD	Bureau van Dijk identifier. Original name: bvdid
DE_BAKISN_CD	Borrower ID ("Nehmernummer")
DE_DESTATIS_CD_STBL	Destatis business register ID. Original name: WE_ID_ALT
ENTTY_RIAD_CD	RIAD identifier.
LEI	Legal entity identifier
USTAN_CD	USTAN identifier. Original name: ukn
JANIS_CD	JANIS identifier. Original name: poolid
SCHUFA_ID	SCHUFA identifier. Original name: SCHUFA_UNTERNEHMEN_ID
anon_BVD_CD	BVD identifier (anonymized)
anon_ENTTY_RIAD_CD	RIAD identifier (anonymized)
anon_LEI	LEI (anonymized)
anon_USTAN_CD	USTAN identifier (anonymized)
anehmernr	Borrower ID, anonymized ("Nehmernummer, anonymisiert")
anon_SCHUFA_ID	Schufa identifier (anonymized)

### 2.2 Details of Variables

**AWMUS\_CD:** AWMuS identifier. Anonymous. Original name: MLD\_NR

Notes	The Bundesbank database AWMuS is the masterdatabase for all foreign statistics related master and metadata in the Deutsche Bundesbank. Source of master data for the research datasets MiDi, SITS and SIFCT. Note that AWMUS_CD does not include the trailing check-digit included in the original MLD_NR.
Available from – to	1981 – 2023
Type of variable	String with up to 8 digits.
Anonymization	Bundesbank internal anonymous ID, may be used in research datasets.

**BBK\_HM\_CO\_ID1:** JANIS identifier (anonymized)

Notes	The Bundesbank database JANIS contains annual financial statements of German non-financial corporations. Successor to USTAN. Since this anonymization of the JANIS-ID is provided by the department responsible for JANIS, the "BBK_HM_CO_ID1" is not necessarily produced from the same data version, therefore there may not always be a "BBK_HM_CO_ID1"-value for every "USTAN_PLUS_CD"-value in the JANIS-Data available to the RDSC.
Available from – to	1997 – 2021
Type of variable	String with 10 digits, corresponds to the anonymized POOLID.

...

Anonymization anonymized USTANPLUS\_CD (anonymization is done by the data providing Bundesbank Department)

---

**BBK\_HM\_CO\_ID2** : JANIKA identifier (anonymized)

---

Notes The Bundesbank database JANIKA contains annual financial statements of German non-financial corporations that stem from commercial data providers. It is a non-anonymised subsample of JANIS without the company data from the credit rating of the Deutsche Bundesbank.

Available from – to 1997 – 2021

Type of variable String with 10 digits, corresponds to the anonymized POOLID.

Anonymization anonymized USTANPLUS\_CD (anonymization is done by the data providing Bundesbank Department)

---

**BVD\_CD** : Bureau van Dijk identifier. Original name: bvdid

---

Notes The Bureau van Dijk identifier is included in the data by the dataprovider Burau van Dijk. A part of the entries with this ID stems from BvD-Master data acquired directly from BvD. Since the BvD-ID is (for German companies) derived from the ID by the German data provider "Creditreform" (by adding the prefix "DE" to the ID-value), it can be complemented by master datasets from Creditreform. In the case of the data used by the RDSC, data from the "Mannheimer Unternehmenspanel" (MUP), from the Zentrum für Europäische Wirtschaftsforschung (ZEW), which is derived from Creditreform data, is used to complement the available BvD-Data. For this purpose, the Creditreform ID is transformed into the BvD-ID by adding the prefix "DE" to the ID values. In consequence, in the data that enters our record linkage, entries that stem from BvD as well as entries that stem from the MUP are identified by this ID.

Available from – to 2004 – 2023

Type of variable String with up to 20 characters.

Anonymization ID may not be used in anonymous datasets. For an anonymized version of this ID see anon\_BVD\_CD

---

**DE\_BAKISN\_CD** : Borrower ID ("Nehmernummer")

---

Notes The Bundesbank database "BAKIS-M" holds bank supervision reference data on borrowers. It contains master data on all borrower entities with a large credit satisfying the reporting requirements to the Deutsche Bundesbank as defined in the KWG. Apart from the borrower-lender level master data it also contains information on their credit of 1 Million or more. BAKIS-M is the source of master data for analytical and research datasets generated from BAKIS-M.

Available from – to 2002 – 2018

Type of variable String of 7 or 8 digits, format as defined by the RIAD data standard.

...

...

Anonymization	ID may not be used in anonymous datasets. There is no anonymized standard version of this ID.
---------------	---

---

**DE\_DESTATIS\_CD\_STBL**: Destatis business register ID. Original name: WE\_ID\_ALT

---

Notes	The official business register ("URS") of the German Statistical Office ("DESTATIS") uses this ID.
Available from – to	2012 – 2021
Type of variable	String of 9 digits as defined by DESTATIS.
Anonymization	ID may not be used in anonymous datasets. There is no anonymized version of this ID.

---

**ENTTY\_RIAD\_CD**: RIAD identifier.

---

Notes	The Bundesbank database "BBk-RIAD" (RIAD: "Register of Institutions and Affiliates Database") is the Bundesbanks' master database for AnaCredit (german part).
Available from – to	2019 – 2023
Type of variable	String with up to 50 characters, format as defined by the RIAD data standard.
Anonymization	ID may not be used in anonymous datasets. There is no anonymized version of this ID yet, since there is no research dataset for AnaCredit yet.

---

**LEI**: Legal entity identifier

---

Notes	The LEI-data from the Global Legal Entity Identifier Foundation issues and maintains this identifier.
Available from – to	2018 – 2023
Type of variable	String with up to 20 characters as defined by the Local Operating Units of the Global Legal Entity Identifier System (GLEIS).
Anonymization	ID may not be used in anonymous datasets. For an anonymized version of this ID see anon_LEI.

---

**USTAN\_CD**: USTAN identifier. Original name: ukn

---

Notes	The Bundesbank database USTAN is a repository that includes master data on companies that have been reported to the Deutsche Bundesbank in the context of its refinancing operations and later for credit assessment purposes. Apart from master data, it contains HGB and IFRS annual financial statements for companies, insolvency data, data reported for the credit register and rating information. Source of master data for USTAN. For our record linkage, the master data in USTAN, which is mostly limited to recent years, is complemented by historical (historized) master data from the original source databases of USTAN ("Jalys" and "Cops").
-------	--

...

...

Available from – to	1987 – 2018
Type of variable	String with up to 8 digits.
Anonymization	ID may not be used in anonymous datasets. For an anonymized version of this ID see anon_USTAN_CD.

---

**JANIS\_CD:** JANIS identifier. Original name: poolid

Notes	The Bundesbank database JANIS contains annual financial statements of German non-financial corporations, including public data sources. Successor to USTAN.
Available from – to	1997 – 2021
Type of variable	String with 10 digits, corresponds to the anonymized POOLID.
Anonymization	ID may not be used in anonymous datasets. For an anonymized version of this ID see BBK_H_CO_ID1.

---

**SCHUFA\_ID:** SCHUFA identifier. Original name: SCHUFA\_UNTERNEHMEN\_ID

Notes	Financial statement data of non-financial companies (only individual accounts) from Schufa Holding AG.
Available from – to	2015 – 2022
Type of variable	String with up to 8 digits.
Anonymization	ID may not be used in anonymous datasets. For an anonymized version of this ID see anon_SCHUFA_ID.

---

**anon\_BVD\_CD:** BVD identifier (anonymized)

Notes	Anonymized version of BVD_CD.
Available from – to	2004 – 2023
Type of variable	String with 64 characters.
Anonymization	anonymized (SHA256 hash based on BVD_CD).

---

**anon\_ENTTY\_RIAD\_CD:** RIAD identifier (anonymized)

Notes	Anonymized version of ENTTY_RIAD_CD
Available from – to	2019 – 2023
Type of variable	String with 64 characters.
Anonymization	anonymized (SHA256 hash based on ENTTY_RIAD_CD).

---

**anon\_LEI:** LEI (anonymized)

Notes	Anonymized version of LEI.
Available from – to	2018 – 2023
Type of variable	String with 64 characters.

...

...

Anonymization      anonymized (SHA256 hash based on LEI).

---

**anon\_USTAN\_CD:** USTAN identifier (anonymized)

---

Notes                      Anonymized version of USTAN\_CD.  
Available from – to      1987 – 2018  
Type of variable         String with 64 characters.  
Anonymization         anonymized (SHA256 hash based on USTAN\_CD).

---

**anehmernr:** Borrower ID, anonymized (“Nehmernummer, anonymisiert”)

---

Notes                      Anonymized version of DE\_BAKISN\_CD.  
Available from – to      2002 – 2018  
Type of variable         String with 8 characters.  
Anonymization         Project specific anonymization procedure.

---

**anon\_SCHUFA\_ID:** Schufa identifier (anonymized)

---

Notes                      Anonymized version of SCHUFA\_ID.  
Available from – to      2015 – 2022  
Type of variable         String with 64 characters.  
Anonymization         anonymized (SHA256 hash based on SCHUFA\_ID).

---

## References

- Bundesamt für Sicherheit in der Informationstechnik. (2020). *Kryptographische Verfahren: Empfehlungen und Schlüssellängen*. BSI TR-02102-1, 2020-01.
- Gabor-Toth, E., Schild, C.-J., and Walter, S. (2023a). *Linking Deutsche Bundesbank Company Data* [Technical Report 2023-05].
- Gabor-Toth, E., Schild, C.-J., and Walter, S. (2023b). *Understanding Overlaps between Different Company Data* [Technical Report 2023-06].