

Data Observer

Ann Hipp, Michael Fritsch*, Maria Greve, Jutta Günther,
Marcel Lange, Christian Liutik, Beate Pfeifer,
Mariia Shkolnykova and Michael Wyrwich

Comprehensive Patent Data of the German Democratic Republic 1949–1990

<https://doi.org/10.1515/jbnst-2022-0058>

Received September 20, 2022; accepted September 26, 2022

Abstract: We document the generation and the content of the Comprehensive Patent Database (CPDB) of the German Democratic Republic (GDR) (1949–1990), which is freely available at GESIS <https://doi.org/10.7802/2448>. The database contains all patents granted in the GDR and published by the Office of Inventions and Patents (later: German Patent and Trade Mark Office/DPMA). The core database covers the years 1950–1990 and contains 24 variables with manually cleaned and processed information on a total of 261,822 unique patents of the GDR.

Keywords: patent data, German Democratic Republic (GDR), invention

JEL Classification: O31, O33, P29, P39, B24

1 Purpose

Patent data are a key source for analyzing innovation activity (Acs et al. 2002). Accordingly, patents play a key role in most studies of innovation and

***Corresponding author: Michael Fritsch**, Friedrich Schiller University Jena, Jena, Germany, E-mail: m.fritsch@uni-jena.de. <https://orcid.org/0000-0003-0337-4182>

Ann Hipp, Jutta Günther, Marcel Lange and Mariia Shkolnykova, University of Bremen, Bremen, Germany, E-mail: ann.hipp@uni-bremen.de (A. Hipp), jutta.guenther@uni-bremen.de (J. Günther), marcel.lange@uni-bremen.de (M. Lange), mariia.shkolnykova@uni-bremen.de (M. Shkolnykova)

Maria Greve, Utrecht University, Utrecht, The Netherlands; and Friedrich Schiller University Jena, Jena, Germany, E-mail: m.greve@uu.nl

Christian Liutik and Beate Pfeifer, Friedrich Schiller University Jena, Jena, Germany, E-mail: christian.liutik@uni-jena.de (C. Liutik), beate.pfeifer@uni-jena.de (B. Pfeifer)

Michael Wyrwich, University of Groningen, Groningen, The Netherlands; and Friedrich Schiller University Jena, Jena, Germany, E-mail: m.wyrwich@rug.nl

technological developments in the communist East German State, the German Democratic Republic (GDR), which existed from 1949 till 1990.¹ However, the patent data used in these studies are often rather incomplete,² and it is unclear to what extent these data have been subject to plausibility checks and data cleaning.

In order to provide a solid empirical basis for further analyses of innovation activities in the GDR, we compiled a Comprehensive Patent Database (CPDB) as part of a collaborative research project funded by the Federal Ministry of Education and Research in Germany (2018–2023).³ The database is freely accessible at GESIS, Leibniz Institute for Social Sciences (<https://doi.org/10.7802/2448>). Compared to other sources of GDR patents, our database has three main advantages. First, it is much more comprehensive than any other available source.⁴ Second, erroneously detected data was manually verified and, if necessary, corrected or completed.⁵ Third, the data contains information that allows for a meaningful assignment of inventive activity to certain locations at the organizational level (for details, see Section 6).

We begin this technical report and overview with a brief description of the GDR patent system (Section 2). Section 3 informs about the methods applied in gathering and preparing the data and Section 4 describes the data in some detail. Finally, we describe the generation of regional identifiers for 17,772 patents of the years 1989 and 1990 (Section 5).

2 Patenting in the GDR

After the foundation of the GDR, a patent law based on the Soviet-type model was established (Wiessner 2015). Based on the GDR Patent Act of 1950, the Office for

1 For example, Bähr and Petzina (1996; 1999), Grupp et al. (2002), Glitz and Meyersson (2020).

2 Glitz and Meyersson (2020) for example used aggregate data of the number of patent applications (not approved patents) at the level of combines (Kombinate) that is limited to the period 1970–1989.

3 Project title “Modernisierungsblockaden in Wirtschaft und Wissenschaft der DDR” (Obstacles to Modernization in the Economy and Science of the GDR) (project number 01UJ1806DY).

4 The recently published Patentcity data (Bergeaud and Verluise 2022) covers only slightly above half of the patents in our database. The coverage of Patentcity varies considerably between single years and sub-periods. Patentcity records considerably lower numbers of patents between the mid-1960s and the late 1970s. Moreover, in more than 80% of the observations, the Patentcity data does not contain the location of the inventor or of the applicant that was published in the years 1989 and 1990. If locations are included, they are quite often rather imprecise.

5 We found a number of typing errors and multiple counts in the original data. Moreover, there were some handwritten annotations on some patent documents that we included in the dataset if relevant. Quite frequently, some International Patent Classification (IPC) codes were added to the original patent ex-post, which we included in the data.

Inventions and Patents (AfEP) in Berlin, subordinated to the Ministry of Planning, was responsible for all issues connected with patenting. The GDR joined the World Intellectual Property Organization (WIPO) in 1968, only one year after the WIPO's foundation. This ensured uniform international standards for filing patents—i.e., the degree of novelty, the inventive level, and the technical applicability (WIPO 1970). Despite the similarities in the dynamics of filing and granting patents, the comparability of the GDR patent data with those of the Federal Republic of Germany (FRG) remains highly controversial (see Fritsch et al. 2022a; Günther et al. 2020a; Grupp et al. 2005; Kogut and Zander 2000).

There were two types of patents in the GDR: the “exclusive patent” (*Ausschließungspatent*) and the “economic patent” (*Wirtschaftspatent*). The exclusive patent guaranteed the patent holder the right to exploit the invention and was very similar to the FRG patent. This type of patent was mainly effective for foreign applicants, especially from Western countries, who were granted exclusive rights for 18 years to use the registered invention in the GDR (Wiessner 2015). The dominant type of GDR patent was the “economic patent”, which was assigned to the respective inventor but did not grant him or her any exploitation rights for the invention; the right of use was instead transferred to the entire economy of the GDR, and the usage was allowed after the responsible central authority had been notified. All inventors who successfully applied for a patent as employees of a state-owned enterprise (VEB), research institute, or other public institution received an economic patent, which recognized him or her as the intellectual source. In case that the patented invention generated economic benefits, the inventor received some remuneration (Wiessner 2015).

3 Methods

3.1 Data Source

With the unification of the two German states in 1990 and the Unification Treaty, the patent law of the FRG became effective in the territory of the former GDR. The patents of the AfEP were transferred to the German Patent and Trademark Office (DPMA) and included in the German Patent Information System (DEPATIS) database (DPMA 2019a). For the compilation of the CPDB, we used the online version of the document archive *DEPATISnet*. DEPATIS contains data on granted patent applications of the GDR for the period from 1949 to 1990, including both economic and exclusive patents (DPMA 2021a, 2021b). The *DPMAregister* platform, which is the official register for all patent applications, also provides an electronic data set with current legal and procedural status information from 1981 onwards (DPMA

2019b). However, considering the information available, this data set is incomplete, as not all data from the GDR patent specifications were correctly transferred to DEPATIS.

3.2 Data Preparation

All corrections and extensions performed for the published CPDB concern the inventor and applicant data as well as the application and publication dates. In case the data extraction performed by the DPMA in the process of patent digitization resulted in incorrect data, we sought to correct all detected errors manually.

The following data preparation steps were performed:

Step 1: The raw data was extracted from the *DEPATISnet* database by the Patent Information Centre of the Friedrich Schiller University Jena in the year 2019 based on a search query. Data access was limited to a maximum of 1,000 patents per query. The project team at the Friedrich Schiller University Jena processed the raw data and created the database.

Step 2: The team at the University of Bremen was responsible for further processing of these data. This particularly included manual data cleaning as well as subsequent creation, preparation, and archiving of the database.

Step 3: The team at the Friedrich Schiller University Jena created an additional module with regional information for the patents of the years 1989 and 1990 (see Section 5).

Each summary table downloaded in step 1 contained the following information listed in columns: (1) serial number, (2) publication number, (3) filing date, (4) publication date, (5) IPC main class, (6) IPC minor/index classes, (7) reclassified IPC, (8) Prüfstoff IPC, (9) inventor, (10) applicant/owner, (11) title, (12) abstract, (13) PDF URL, and (14) searchable text URL. This resulted in a total of approximately 300 Excel tables, which were transferred to a database. In addition to the overview tables in Excel format, all related patent specifications were downloaded in PDF format.

The content of columns 1)-6) as well as 9)-12) was automatically retrieved from the patent specification of the DPMA and corresponds to the INID codes of the same names (see Chapter 3). Columns 7)-8) and 13)-14) reflect information that is not directly contained in the patent specification. Column 7) stands for the state-of-the-art adapted and reclassified IPC classes. Since the technical definitions may change with each IPC revision, individual documents were reclassified using WIPO's Master

Classification Database (MCD). The reclassification is not done in the patent document but in the database (DPMA 2020). Column 8) entails an adjusted technological classification. This adjustment was made by the experts of the DPMA, who believe that in certain cases the technological area specified in the patent document is not accurately and/or solely reflected in the given IPC code(s). This column may contain not only internationally listed IPC codes, but also codes from a more refined German patent classification (“*Deutsche Patent-Feinklassifikation*”, DEKLA). These so-called Prüfstoff IPC and reclassified IPC are not assigned in all cases by the DPMA examining staff, and therefore, columns 8) and 9) have many missing values that cannot be filled out manually. Columns 13) and 14) represent hyperlinks that contain the copy of the original patent specification and the patent text that has been generated from the original document by using automatic Optical Character Recognition (OCR) conversion, respectively. The latter document is deprived of any drawings and allows a quick information search within the patent document.

In step 2, the details of the applicants and inventors in all empty fields were manually compared with the data in the respective patent specifications, digitized, and supplemented. For the applicant/owner, all persons indicated under INID codes (71) and (73) were added from the patent specification. In some cases, the information on the representative (74) was added.

In some cases, the machine reading of the application and publication data by the DPMA led to inadequate recording of the dates. For example, there was an indication that a patent was published in 1968 but not filed until 1988. In such cases, the year of filing and the year of publication were manually checked and corrected. In addition, patents for which the publication date corresponds to the application date were manually checked and the information was corrected.

The date format was adjusted to the standard DD.MM.YYYY for all patents.

The placeholders in the title and abstract were removed and the language information from the text (DE or EN) was transferred to a new column (sprachetitel and spracheabstract). Obvious transfer errors that led to incorrect information in the title were corrected manually. The CPDB was prepared according to the formal requirements of the Leibniz Institute for the Social Sciences (GESIS) and the necessary information was added.

There were several erroneous entries in the database. The respective patent numbers are summarized in Hipp et al. (2022, Table A1).

4 Comprehensive Patent Database (CPDB)

The final database contains all granted patents of the GDR published by AfEP (later DPMA) in the period between January 1, 1939 (application before but granted in the

GDR) and June 29, 2006 (application in but granted after the GDR). The core database covers the years 1950–1990 and contains 24 variables with manually cleaned and processed information on a total of 286,478 patent entries, which correspond to 261,822 unique GDR patents. This difference can be explained by the fact that for around 24,500 patents (around 9% of all entries), several (up to 3) documents of the patent family are included in the database. The CPDB of the GDR is freely available in Excel format (.xlsx) as well as delimited text file (.csv) at the Leibniz Institute for the Social Sciences (GESIS); <https://doi.org/10.7802/2423>.

Information is missing in the database if it is not included in the patent specifications. In several cases, this concerns the IPC secondary classes and the abstract. For the years before 1975, only a few abstracts are available. In total, abstracts are missing for 165,740 patents.

Despite its availability in the patent documents, data may be missing in the database in case of more than one applicant or inventor, because the focus in preparing the data was on completing the empty lines. In addition, it is often not possible to distinguish between applicant and owner if, for example, only headquarters or the patent attorney are named. The patent attorney is included in the variable *anmelderinhaber*, although he or she is not the inventor or applicant. Rarely, there is no information on the inventor in the patent specifications (coded as “wird nicht genannt”) because he or she did not want to be mentioned as an inventor. For the supplementary module with the regional data for 1989/90, the inventor names are considered sensitive data and are anonymized by the designation “inventor”.

In a few cases, there are transcription errors, with the correct data being contained in the patent specifications but not in the database. This is particularly the case for handwritten annotations or special fonts that could not be processed without error by the text-recognition system. The correct information could then be found in the PDF files of the patent specification. In some cases, no PDF URL to the patent specifications could be retrieved; in this case, the searchable text had to be used. Table 1 provides an overview of the variables, their description, and further notes.

5 Regional Information

For a total of 17,772 patents from 1989 to 1990, the CPDB contains additional information that could be usefully recorded because of a change in the database. In the course of a reform of the GDR patent law in June 1990, the documentation of applicants and inventor information was also improved. Whereas previously only the combine or the parent company was mentioned as the applicant, from summer

Table 1: Variable description.

Variable	Description	Document/Notes
veroeffentlichungsnummer	Publication number of the granted patent application assigned by the German Patent and Trademark Office (DPMA)	DPMA INID-Code ^a 11 Note: Corresponds to the columns <code>publn_auth</code> , <code>publn_nr</code> , and <code>publn_kind</code> from the PAT-STAT table <code>tls211_pat_publn</code>
anmeldedatum	Patent application date	DPMA INID-Code 22 (DPMA 2014)
veroeffentlichungsdatum	Date of publication of the patent	DPMA INID-Code 44 or 45 (DPMA 2014)
ipchauptklasse	Main class or technical focus of the patent	DPMA INID-Code 51 (marked bold) Note: PATSTAT does not distinguish between the 1st and other IPC classes
ipcnebenindexklassen	Further technical fields (IPC secondary classes) to which the patent can be assigned	DPMA INID-Code 51 Note: PATSTAT does not distinguish between the 1st and other IPC classes (DPMA 2020)
reklassifizierteipcmcd	State-of-the-art adapted and reclassified IPC classes	Since the technical definitions may change with each IPC revision, individual documents are reclassified using WIPO's Master Classification Database (MCD). The reclassification is not done in the patent document but in the database (DPMA 2020).
pruefstoffipc	More refined German patent classification 'Prüfstoff IPC' according to DEKLA ("Deutsche Patent-Feinklassifikation")	Classification by experts of the DPMA after allocation of the bibliographic IPC (DPMA 2020)
erfinder	Inventor of the patent	DPMA INID-Code 72 (DPMA 2014)
anmelderinhaber	Patent applicant OR/AND patent owner OR/AND representative (usually the patent attorney)	DPMA INID code 71 OR/AND 73 AND/OR 74 (DPMA 2014)
titel	Patent title	INID-Code 54
sprachetitel	Language of the title of the patent	
abstract	Description of the patent	INID-Code 57
spracheabstract	Language of the abstract of the patent	
pdfurl	Link to the patent specification in the DPMA database	

Table 1: (continued)

Variable	Description	Document/Notes
recherchierbarertexturl	Machine-readable document of OCR without graphics	May be incorrect because automatically generated
nameapplicantcleaned	Detailed recording of the name of the patent holder	Only available for the years 1989 and 1990 (see Section “Regional information”). The inventors’ names are anonymized.
type	Type of patent holder (1: foreign country; 2: GDR company; 3: GDR inventor; 4: foreign country with participation of GDR inventors)	Only available for the years 1989 and 1990 (see Section “Regional information”)
gemeinde_ ^b	Municipality number (as of 2021) where the seat of the patent holder was located	Only available for the years 1989 and 1990 (see Section “Regional information”)

^aINID codes are numbers for identifying bibliographic data according to WIPO standard ST. 9 and are used by all patent offices. From 1977 or DD123500, the INID codes are noted on the GDR patents. ^bNumbers 1 to 7.

of 1990 onwards, the actual applicant establishment of the combine was also named in the entries. Furthermore, some of the address data of the inventor(s) was entered. In practice, this change means that GDR patent applications of the years 1989 and 1990, which were granted from summer 1990 onwards, either before or after German reunification, open up the possibility to assign a patent to a certain location, thereby identifying the regional structure of innovation activities in the GDR. For patent applications prior to 1989, we know only the headquarter of the respective combine (*Kombinatszentrale*) that is indicated as applicant in the patent specification. Since the establishment of the combine that generated the invention was often located in considerable distance to the headquarter, there was no way to link a patent to a certain region. Moreover, for the indicated period it is also possible to distinguish the exact location of the research institute.⁶

Against this background, we additionally extracted information from the original PDFs of the patent specifications for the GDR patents of the years 1989 and 1990 in order to identify the location of the patent applicants/owners in more detail

⁶ While in the earlier years, the location is mostly allocated to the Academy of Science in Berlin, the regional module allows identifying the precise location. It turns out that among all patents registered by the Academy of Science, roughly half of them stem from Berlin, while the others were scattered around the GDR. Overall, less than 10% of all patents in 1989 and 1990 were applied for by the Academy of Science.

than in the DPMA database. After cleaning the applicant and owner data, we assigned a municipality (*Gemeinde*) number (as of 2021) to each applicant. For patents with several applicants in different municipalities, several municipality numbers (up to 7 municipalities) were assigned to the patent. If the patent applicants are also the inventors, they are assigned a municipality number based on the inventors' place of residence. Patents of foreign applicants are marked separately.

References

- Acs, Z.J., Anselin, L., and Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Res. Pol.* 31: 1069–1085.
- Baar, L. and Petzina, D. (1999). *Deutsch-deutsche Wirtschaft 1945 bis 1990: Strukturveränderungen, Innovationen und regionaler Wandel. Ein Vergleich.* Scripta Mercaturae, St. Katharinen.
- Bähr, J. and Petzina, D. (Eds.) (1996). *Innovationsverhalten und Entscheidungsstrukturen.* Duncker & Humblot, Berlin.
- Bergeaud, A. and Verluise, C. (2022). A new dataset to study a century of innovation in Europe and in the US. Discussion Paper no. 1850, Centre for Economic Performance. London School of Economics and Political Science.
- DPMA (2014). *DPMAinformativ. Normierung von Patentliteratur.* Deutsches Patent- und Markenamt, Munich.
- DPMA (2019a). *1986 to 2000: the German Patent Office until the turn of the millennium.* Deutsches Patent- und Markenamt, Munich.
- DPMA (2019b). *DPMA register. Dateninhalt.* Deutsches Patent- und Markenamt, Munich.
- DPMA (2020). *DPMAinformativ. Patentrecherche mit Klassifikationssymbolen.* Deutsches Patent- und Markenamt, Munich.
- DPMA (2021a). *30 Jahre Deutsche Einheit – zwei deutsche Patentämter wurden eins.* Deutsches Patent- und Markenamt, Munich, Available at: https://www.dpma.de/dpma/wir_ueber_uns/geschichte/30jahre_einheit/index.html.
- DPMA (2021b). *Datenbestand.* Deutsches Patent- und Markenamt, Munich, Available at: <https://depatisnet.dpma.de/DepatisNet/depatisnet?window=1&space=menu&content=index&action=datenbestand>.
- Fritsch, M., Greve, M., and Wyrwich, M. (2022). Shades of a socialist legacy? Innovation activity in East and West Germany 1925–2014. *Jena Economic Research Papers #2022–12*, Friedrich Schiller University Jena.
- Glitz, A. and Meyersson, E. (2020). Industrial espionage and productivity. *Am. Econ. Rev.* 110: 1055–1103.
- Grupp, H., Dominguez-Lacasa, I., and Friedrich-Nishio, M. (2002). *Das deutsche Innovationssystem seit der Reichsgründung.* Physica-Springer, Heidelberg.
- Grupp, H., Dominguez Lacasa, I., Friedrich-Nishio, M., and Jungmittag, A. (2005). Innovation and growth in Germany over the past 150 years. In: Cantner, U., Dinopoulos, E., and Lanzillotti, R.F. (Eds.), *Entrepreneurships, the new economy and public policy.* Springer, Berlin/Heidelberg, pp. 267–287.

- Günther, J., Hipp, A., and Ludwig, U. (2020). *Universalien der Innovation – Erfindertum und technischer Fortschritt in der DDR und dessen Rolle nach 1990*. University of Bremen: Institute for Economic Research and Policy. <https://doi.org/10.26092/ELIB/332>.
- Hipp, A., Fritsch, M., Greve, M., Günther, J., Lange, M., Liutik, C., Pfeifer, B., Shkolnykova, M., and Wyrwich, M. (2022). Comprehensive patent data of the German Democratic Republic 1949–1990—technical report and dataset overview. Jena Economic Research Papers #2022-011, Friedrich Schiller University Jena.
- Kogut, B. and Zander, U. (2000). Did socialism fail to innovate? A natural experiment of the two Zeiss companies. *Am. Socio. Rev.* 65: 169–190.
- Wiessner, M. (2015). Das Patentrecht der DDR. In: Otto, M. (Ed.), *Geschichte des deutschen Patentrechts, Geistiges Eigentum und Wettbewerbsrecht*. Mohr-Siebeck, Tübingen, pp. 239–288.
- WIPO (1970). Paris convention for the protection of industrial property, Paris Notification No. 16. Entry into force of Articles 1 to 12 of the Stockholm Act.

Publications based on the data

- Fritsch, M., Greve, M., and Wyrwich, M. (2022). Shades of a Socialist Legacy? Innovation Activity in East and West Germany 1925–2014. Jena Economic Research Papers #2022–12, Friedrich Schiller University Jena.
- Günther, J., Hipp, A., and Ludwig, U. (2020). *Universalien der Innovation – Erfindertum und technischer Fortschritt in der DDR und dessen Rolle nach 1990*. University of Bremen: Institute for Economic Research and Policy. <https://doi.org/10.26092/ELIB/332>.
- Hipp, A., Ludwig, U., and Günther, J. (2021). Unable to innovate or just bad circumstances? Comparing a state-led and market-based innovation system in Germany. *Papers on Economics & Innovation* #2111. University of Bremen, <https://doi.org/10.26092/elib/1330 >.
- Hipp, A., Ludwig, U., and Günther, J. (2023). Und sie wirkten doch: Patente und Produktivität in der DDR. In: Flade, F., Steinkamp, A., and Walerski, K. (Eds.), *Transformation in Polen und Ostdeutschland: Neue Einblicke in die Debatte über eine gesplattene Wirtschaft*. Harrassowitz Verlag, Wiesbaden (forthcoming).