Newspapers in German Language – Progress in Digitisation and Accessibility

Ulrich Hagenah
Regional Library Department, Digitisation Unit, State and University Library, Hamburg, Germany
E-mail address: hagenah@sub.uni-hamburg.de

Copyright © 2015 by Ulrich Hagenah. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License: http://creativecommons.org/licenses/by/3.0/

Abstract:

Newspapers have been more and more recognised as key sources of modern history. The 100 years anniversary of World War I showed this in diverse examples of online presentations, publications and exhibitions – and it gave a heavy boost to academic demand as well as digitisation efforts for newspapers in German language.

Federal political structures and insufficient funding policies have been two of the reasons why the state of German newspaper digitisation is still comparatively delayed and dissatisfactory. But there are a couple of developments going on which show improvements in quantity and quality. The paper tries to summarize the most important ones focusing on topics like:

- the digitisation efforts of several regional libraries
- the results of the Europeana Newspapers Project regarding newspapers in German language
- the piloting project of the German Research Foundation with prototype digitisation projects in five libraries, with the complete digitisation of German newspapers from the 17th century amongst others
- the adaption of new features in order to optimise the presentation of newspapers in the German Union Catalogue of Serials

A gradual change in digitisation strategies can be acknowledged in Germany, driven by academic demand as well as the stewardship role of specifically interested libraries. With a short side glance at the digitised German-language press outside Germany the paper concludes the state of accessibility and steps still to be undertaken on the way to a comprehensive or at least representative usage of historical newspapers in German.

Keywords: newspaper, digitisation, Germany, historical press, German Union Catalogue of Serials
0 Introduction

In the last two decades newspapers have been growingly recognised as key sources of modern history. The 100 years anniversary of World War I showed this in diverse examples of online presentations, publications and exhibitions – and it gave a heavy boost to academic demand as well as digitisation efforts for newspapers in German.\(^1\) We also observe that academic researchers in the history of our language and literature, of social ideas and political ideologies more and more often resort to newspaper text as a resource for text and data mining over the decades and centuries. The Deutsches Textarchiv\(^2\) may serve as an example: it is a common research project of the consortium of German academies. A representative sample of text corpora in German across all genres and disciplines from 1600 up to the 1990s is being established in order to provide a reliable statistical basis for historical linguistic research. Researchers of the German academies firstly formed a comprehensive bibliography as groundwork for the selection of text samples to be digitised. The presentation of results displays those for newspaper text separately, comparable to scholarly texts, functional writing and fiction. All texts are downloadable and the statistical analysis reusable for projects of content analysis or discourse analysis.

---


\(^2\) http://www.deutschestextarchiv.de/ [17.7.2015].

---

Deutsches Textarchiv, Statistics for the term „Revolution“, according to text genres, 1600 – 1990, and with display of hits for single text corpora
Preconditions for matching this growing demand for digital newspapers were not too good in Germany. As a federal republic consisting of 16 federal states Germany does not have available central institutions leading the cultural and scientific sectors - the federal states bear responsibility for culture and science, in the field of librarianship for the regional libraries and of course university and other academic libraries. The German National Library, founded in 1913 as Deutsche Bücherei in Leipzig, being a comparably young institution without historical holdings dating further back, exerts a leading and coordinating role in most of the current librarian affairs, but mostly does not so in the field of retrodigitisation of historical material.

So, generally spoken, there is no institution at national level responsible for steering and coordinating the digitisation strategies. We are used to starting regional initiatives which quite often tend to be followed by others or even lead to contracting and common projects at national level. Another actor in the academic field and most important for the funding of research-oriented library services is the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). The DFG has initiated several funding schemes in respect of digitisation, for incunabula, for 17th century and 18th century monographs and journals for instance, for joint research projects of academic institutions and libraries, and for several other purposes. But it took much more time until the DFG decided to enter into the field of newspaper digitisation, longer than libraries could wait and postpone activities. I will come back to this later, first steps have been done by many libraries in the past few years.

And that is what I would like to report in a short overview: diverse activities and results of newspaper digitisation over the past couple of years\textsuperscript{3}, in five steps and with some concluding remarks on future prospects.

1 Regional libraries

Each of the 16 German states has one or more regional libraries, responsible for collecting, maintaining and promoting the cultural heritage of their respective region in cooperation with archives, museums and other heritage institutions. And many of the bigger university libraries also play an important role in making their regional and special collections digitally accessible as far as they are relevant for research and for the broader public – which definitely applies to their newspaper heritage.

A rough survey conducted in April 2015 shows:

Regional and university libraries in Germany⁴ have up to now digitised at least 633 newspaper titles comprising 7,874 annual volumes. 62 % of these digitised newspapers were published before 1900. The amount of pages per annum widely differs from 1.000 - 1.500 for smaller provincial papers published less than six times a week to 3.500 - 4.500 for the metropolitan newspapers of the decades from 1880 onwards. Supposed an average of 1.500 pages per year, we have a total of 11,8 million pages of almost 650 newspaper titles digitised up to April 2015. For 50 of these newspapers more than 50 annual volumes have been digitised. Many of the long-running and most of the big-city papers, party papers and opinion leading media are still missing in this summation. And the outcome of two special projects in Munich and Bremen has to be added to these figures, I will come back to them very soon.

Certain disadvantages of the decentralized approach are obvious: Standards applied in the digitisation processes vary from one to the other state or even library; the libraries themselves decide about their respective software platforms. OCR is an established standard in many, mainly the larger digitisation projects, but not in all of them. Resources are slender, coordination lacking very often. Presentation interfaces are individual, about 30 different ones are on offer at the moment. On the other hand, one of the regional libraries’ special assets is their interconnectedness in the region they are responsible for. They are inevitably part of a network of cultural institutions in the area; most of them cooperate and network with all kinds of cultural and heritage institutions. Thereby cooperative efforts for digitisation emerge as a natural consequence as they did a couple of years ago for cooperative microfilming. For example the Bonn University and Regional Library⁵ systematically surveyed the newspaper holdings of the city libraries and archives of its catchment area in order to digitise all newspapers of the former Rhine Province.

**Bonn University and Regional Library, Homepage “Zeitungen”**

---

⁴ Figures except the Bavarian State Library, Munich, and the Bremen State and University Library whose current projects are reported separately below.  
⁵ [http://digitale-sammlungen.ulb.uni-bonn.de/nav/classification/229854](http://digitale-sammlungen.ulb.uni-bonn.de/nav/classification/229854) [17.7.2015].
This is a quite promising effort to cover a whole region because in digitisation matters local archives primarily lie their focus on their unique records, manuscript and document holdings, but not as much on printed material. Similar activities also happen in Thuringia and in the City State of Hamburg. Some states established support programs for the maintenance of important historical sources, encompassing the newspapers of the region. So there are a lot of activities going on without nationwide coordination.

2 Bavarian State Library

The Bavarian State Library, Munich (Bayerische Staatsbibliothek, abbreviated BSB), has to be mentioned separately in respect of newspaper digitisation. The BSB with its currently 9.8 million books is one of Europe’s most important universal libraries and also regional library for the Free State of Bavaria. Besides many famous and voluminous collections of diverse media, its historical stock encompasses one of the foremost German newspaper collections.

The BSB was the first German library to cooperate with Google in a mass digitisation project within the framework of a public-private partnership. The BSB digitised its complete copyright-free historical holdings of printed works over several years, presumably over one million titles. Among these are about 1.000 newspaper titles from Bavaria and also from other parts of Germany up to the year 1900. Many of them are already available on the web. For their presentation the BSB is developing a new interface including calendar and full-text search which will replace the current presentation of newspapers in the Bavarian regional portal Bayerische Landesbibliothek Online (BLO) and Digipress. Besides the ZEFYS newspaper information system of the Berlin State Library and those various smaller collections of historical German press the BSB in Munich will provide one of the foremost access points to historical newspapers in Germany.

3 Europeana Newspaper Project (ENP)

Two German libraries, the Berlin State Library and Hamburg State and University Library, took part in this far-reaching European project, which ended after three years in March 2015. 18 project partners and eleven so called associated partners, most of them national libraries from all over Europe, some regional libraries and service providers, collaboratively built up a European newspaper browser under the aegis of The European Library (TEL) containing more than 3,6 million issues of newspapers from almost 3.500 newspaper titles, with 20,6 million full-text indexed pages. Moreover, there are some 10 more millions of images available without full-text indexing. They cover a timespan from 1618 to 1955. The browser is to be converted into a newspaper channel of the

---

6 For example North-Rhine Westphalia: [http://www.ub.uni-koeln.de/bibliothek/profil/projekte/digirheinschriften/index_ger.html](http://www.ub.uni-koeln.de/bibliothek/profil/projekte/digirheinschriften/index_ger.html) [17.7.2015].
8 Searchable via Zeitschriftendatenbank, [http://dispatch.opac.db-nb.de/DB=1.1/CMD?ACT=SRTCHA&SRT=1ST Ty&TRM=STW+all+20%20VER+20muenchen+DOK+zt+und+Id](http://dispatch.opac.db-nb.de/DB=1.1/CMD?ACT=SRTCHA&SRT=1ST Ty&TRM=STW+all+20%20VER+20muenchen+DOK+zt+und+Id) or the local Online Catalogue, [https://opacplus.bsb-muenchen.de](https://opacplus.bsb-muenchen.de) [17.7.2015].
9 [http://www.bayerische-landesbibliothek-online.de/schwerpunkte#Zeitungen](http://www.bayerische-landesbibliothek-online.de/schwerpunkte#Zeitungen) [17.7.2015].
10 [http://digipress.digitale-sammlungen.de/de/fs1/start/static.html](http://digipress.digitale-sammlungen.de/de/fs1/start/static.html) [17.7.2015].
12 [http://www.europeana-newspapers.eu](http://www.europeana-newspapers.eu) [17.7.2015].
Europeana platform within 2015 or 2016, so that the future access to European newspapers will be contextualised by this vast assemblage of historical sources in many genres and media formats.

Berlin and Hamburg as the foremost centres of mass media production in Germany through the 18th to 20th centuries contribute some long runs of very important newspapers to the Europeana, e.g. the Berliner Tageblatt, Berliner Volkszeitung, Hamburger Nachrichten, Hamburger Anzeiger or the two business papers Berliner Börsenzeitung and Börsen-Halle (Hamburg). Research about World War I exemplified clearly how fruitful the cross-border and transcultural search through newspapers can be. It has become much easier to compare views on the same events from all warring parties and to change perspectives using the leading or the provincial media of different countries simultaneously.

Europeana Newspaper Project. Homepage

1.2 million out of 3.6 million newspaper issues of ENP are in German resp. bi- or trilingual with German being one of the languages besides Russian and Estonian / Latvian in Estonia and Latvia for example, or besides Luxembourgian and French in Luxembourg. By far the largest contribution from outside Germany comes from the well-established Austrian ANNO portal which introduced full-text search into its interface for 1689-1918 and 1938-1944 and is continually expanding. Currently it contains 183 different newspapers in German language, with a total of 4.150 annual volumes, one half each published before 1900 and 1900-1944.

http://www.europeana.eu/portal/ [17.7.2015].
http://anno.onb.ac.at/index.htm and http://www.theeuropeanlibrary.org/tel4/newspapers [17.7.2015]. ANNO does not only contain newspapers, but also about 16.000 play bills from theatres of the Austrian-Hungarian Empire and many professional journals, yearbooks, annual reports, etc. http://anno.onb.ac.at/anno-suche/#searchMode=simple&resultMode=list&from=1; full-text search for 1919-1937 is said to be added soon [20.7.2015].
Another important bundle of ENP newspapers in German derives from Italy, the Landesbibliothek Friedrich Tessmann in Bozen, South Tyrolia. They delivered 30 newspapers with more than 900 annual volumes, out of which more than 60% date from the 20th century. Filtering by language is one of the basic features of the ENP browser – thereby it is easy to identify the German newspapers in all countries taking part.

4 The pilot scheme “Digitisation of Historical Newspapers” of the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG)

As mentioned above, it took several years of discussion, preparation and conceptualising, with a first DFG conference of scholars and librarians in Bremen in 2009 as a starting point, to initiate a pilot scheme funded by the DFG as a preliminary step to set up a nationwide funding scheme for newspaper digitisation. Six project partners conduct prototype digitisation projects, work on adapting the functions of the software platform Goobi and the so-called DFG viewer for newspapers, and finally there is an important sub-project destined to further develop the presentation of newspapers in the German Union Catalogue of Serials. The pilot scheme, running from May 2013 until December 2015 is destined to lead to a master plan for the mass digitisation of newspapers in Germany.

The first and overarching aim of the master plan is to support the digital transformation of historical newspapers as a resource for research and teaching in various disciplines. The sub-projects of the pilot scheme focus on testing different ways and means of digitising, the usage of microfilms, costs and results of various indexing methods, on the implementation of standards and the optimising of tools for the presentation of data and metadata. Also recommendations for selection criteria and funding standards are expected as an outcome of the joint pilot scheme and intensive discussions with academic users such as historians, media scholars, linguists and social scientists.

The libraries in Dresden, Halle and Munich digitised newspapers from the late 18th to the 20th centuries on different levels in their sub-projects, i.e. the Allgemeine Zeitung (1798–1929; Tübingen, http://digital.tessmann.it/tessmannDigital/Zeitungsarchiv/Zeitungen;jsessionid=E33E2E39454B741BC6C6DE1C4B3A7F27 and http://www.theeuropealibrary.org/tel4/newspapers [20.7.2015].

Of course there are some more digitised newspapers in German as well in the ENP presentation (Latvian, Estonian, Slovenian, etc.) as beyond: papers from Switzerland (three titles, 151 annual volumes) and Liechtenstein (nine titles, 239 annual volumes) should be mentioned here, and also the newspapers in Chronicling America with a total of 352 annual volumes, coming from eight American states, most of them from Pennsylvania, cf. http://chroniclingamerica.loc.gov/newspapers/?state=&ethnicity=&language=ger [20.7.2015].

Berlin State Library, Bremen State and University Library, Saxonian Regional and University Library, Dresden, German National Library Frankfurt, Halle University and Regional Library, Bavarian State Library, Munich.


Cf. chapter 5.


Ulm, Augsburg, München), founded by the famous publisher Johann Friedrich Cotta, one of the leading political German newspapers of the 19th century.25

A fourth digitisation sub-project was set up at the Bremen State and University Library where the most extensive collection of German 17th century newspapers resp. reproductions in diverse formats is situated. In the 1970s and 80s the German 17th century press had been registered in a comprehensive bibliography for the first century of newspapers at the Institut Deutsche Presseforschung (University of Bremen).26

Now the complete collection is being digitised and presented as images27: 605 newspaper titles originating from 300 newspapers, containing 60,000 issues and approximately 330,000 pages in total. Copies are added whenever newly found issues emerge anywhere, and the librarians work in close cooperation with several research and editorial projects.

27 http://brema.suub.uni-bremen.de/zeitungen17/periodical/pageview/1076614 ; http://brema.suub.uni-bremen.de/zeitungen17/periodical/pageview/1076540 [20.7.2015].
5 Enhancement of the German Union Catalogue of Serials (Zeitschriftendatenbank, ZDB)

The ZDB\(^{29}\) is „the world’s largest specialized database for serial titles (journals, annuals, newspapers, etc., incl. e-journals). [...] It actually contains more than 1.7 million bibliographic records of serials from the 16\(^{th}\) century onwards - 60.000 newspapers among them -, from all countries, in all languages, held in 4.400 German and Austrian libraries, with 13.6 million holdings information.“\(^{30}\) The ZDB-network is managed by and the maintenance and further development of the ZDB lies with the Berlin State Library and the German National Library. Through its cooperation with the Europeana Newspaper Project the ZDB started expanding its scope of records by including diverse national and European collections of digitised newspapers. Thus it is going to become one of the most important access points for digital and digitised newspapers worldwide.

Inside Germany the Union Catalogue of Serials is a centerpiece of academic infrastructure, and as such it naturally serves as the central knowledge base for coordinating the digitisation process of serials on the whole as well as of newspapers in particular. In respect of newspaper cataloguing the ZDB was often criticised in the past for its presentation of entities. It appeared confusing and was criticised for excessively fragmenting long-running newspapers into small-sized title splits\(^{31}\), which are linked to each other consecutively and with their respective supplements, parallel and local issues, etc., but could not be overlooked comprehensively for the whole timespan of a paper – thus it was quite difficult to get a general idea of a newspaper like the Hamburger Nachrichten, published from 1792-1939, changing its title twice, incorporating another famous paper, the Relations-Courier, having many supplements in the early 20th century and an ‘illegitimate’ successor in the same newspaper title 1998.\(^{32}\)

So the newspaper community highly welcomed the development plan for a new ZDB presentation interface as sub-project of the DFG pilot scheme for historical newspapers. Aims were to thoroughly enhance research and navigation facilities, and to add special features for newspaper search and various visualisation tools in order to facilitate orientation in complex data relationships. The results were published in a beta-version\(^{33}\) at the beginning of July 2015 – and, from my point of view, the new ZDB interface provides an enormous enhancement in quality and user-friendliness.

Some of the new functionalities should be mentioned in brief: A time bar for publication years facilitates the simple or extended search right from the beginning. You can filter results according to facets like media type, country and years of publication, online availability, digitisation projected, language, periodicity, places of distribution, etc. For a single newspaper a graphical overview is given showing the holdings of all libraries possessing issues of the paper (“Bestandsvergleich”).

\(^{29}\) [http://www.zeitschriftendatenbank.de](http://www.zeitschriftendatenbank.de) [20.7.2015].
\(^{30}\) [http://dispatch.opac.dnb.de/DB=1.1/](http://dispatch.opac.dnb.de/DB=1.1/) [20.7.2015].
\(^{31}\) According to the ZETA system of rules ([http://www.zeitschriftendatenbank.de/erschliessung/arbetsunterlagen/zeta/](http://www.zeitschriftendatenbank.de/erschliessung/arbetsunterlagen/zeta/), [20.7.2015]) a new record has to be created for each title that has changed within the first five title words compared to its predecessor, and both are linked to each other. So many long-running newspapers have three, four or more linked title records in the database, and additionally records for all supplements, regional or local editions, etc.
\(^{33}\) [http://beta.zdb-opac.de/zdb/index.xhtml](http://beta.zdb-opac.de/zdb/index.xhtml) [20.7.2015]. It is going to replace the current ZDB version at the end of 2015 as soon as further enhancement measures for all serial media will have been implemented.
You can choose for how many libraries you wish to see the holdings being displayed on a time bar, and you can also choose or exclude certain holdings. It is possible to confine the diagram to certain regions defined for the inter-library loan system. Later also a complete geo-visualisation of the owners will be offered.

Finally, the ZDB proves new ways of clarifying the network structure of a newspaper history via visualisation. So called network graphs represent the whole networks or parts of it – they are scalable ("display more relations", "display less relations") and zoomable. Different colours of the connecting lines represent different kinds of relations (predecessor, successor, supplement, parallel edition, etc.).
Various options for navigation through the network graphs are provided. The chronological sequence of newspaper titles over the whole timespan of the paper can be displayed ("Titelhistorie", "title history"),

as well as an overview over the plenty of supplements and their periods of publication ("Beilagenhistorie", "supplement history").
The graphs also enable the user to jump to the records of related newspapers, issues or supplements and then follow their relational network.

6 Conclusions

Federal political structures and insufficient funding policies have been two of the reasons why the state of newspaper digitisation in Germany is still comparatively delayed and dissatisfactory. But there are a couple of developments going on which show improvements in quantity and quality. Newspaper digitisation is still an uncoordinated decentralized process, and the results still vary in respect of indexing levels, presentation modes and contextualisation. The gradual change in digitisation strategies and many individual initiatives on the regional level were driven by academic demand, multiple requests of the broader public, as well as the stewardship role of specifically interested libraries. The DFG pilot scheme, the Google joint venture of the Bavarian State Library and the success of the Europeana Newspaper Project undoubtedly boosted this process. A funding scheme of the German Research Foundation (DFG) for digitising a representative corpus of newspapers should be the next step in 2016/17.
But even when this will have been accomplished, the absence of a common German portal for historical newspapers will remain a severe deficit – or rather: an objective to work on collaboratively in the future. The most appropriate platform for such a newspaper portal appears to be the German Digital Library (Deutsche Digitale Bibliothek, DDB). “The DDB is Germany’s central, national portal for culture and science. The portal, wholly organized and funded by the public sector, has the ongoing aim of linking up the digitised inventories of Germany’s cultural and scientific institutions and making them available to the public.”

The DDB consists of a network of a constantly increasing number of organisations and institutions (totalling more than 2,100 in March 2014). And it is the national data aggregator for Europeana, funded by the German Federal Government and the German states, with a head office in the Prussian Cultural Heritage Foundation (Stiftung Preußischer Kulturbesitz), Berlin, and the German National Library (Deutsche Nationalbibliothek), Frankfurt, leading the technical coordination department and the Service Center of DDB. Preliminary talks and conceptual thoughts have been initialised regarding an integration of a newspaper platform including full text search into the DDB portal. So further developments in this field may be subject of another progress report on German newspaper digitisation in the future.

---

34 https://www.deutsche-digitale-bibliothek.de/ [20.7.2015].