News Media Section

Visual representations of newspaper information

Jessica Hubrich
German National Library, Frankfurt am Main, Germany.
E-mail address: j.hubrich@dnb.de

Abstract:

The German National Union Catalogue of Serials (Zeitschriftendatenbank – ZDB) provides access to 60,000 newspapers being held in more than 4,000 German and Austrian Libraries. It does not only contain extensive descriptive data about single titles in all languages from the 16th century onwards and corresponding holdings but also about relationships between these titles. Until recently, the ZDB OPAC presented this information exclusively in a traditional text form, focusing on single title records. In 2015, a new ZDB catalogue was created that transcends the traditional OPAC concept. Intuitive search interfaces support users in many ways, interactive graphical representations provide insights into and an overview of interrelations existing between different titles. This paper describes the main visualisations especially relevant for digitisation projects or newspaper research. A beta version of the new ZDB catalogue is available at http://beta.zdb-opac.de/zdb/index.xhtml.

Keywords: German National Union Catalogue of Serials, ZDB, Zeitschriftendatenbank, catalogue, OPAC, search interface, visualisation, newspapers, title relations, title connections, title history, holdings overview

1. The German Union Catalogue of Serials

The German National Union Catalogue of Serials (Zeitschriftendatenbank – ZDB) is one of the world's largest databases for journals, newspapers, monographic series and other serial publications. It is maintained and further developed by the German National Library and the Berlin State Library. As of March 2016, it contains around 1.8 million bibliographic records describing resources kept in about 4,000 German and Austrian libraries. Among these are about 60,000 newspapers. These are not restricted to a specific language, country or time period and may be available in printed, electronic or digitised form. 13.5 million records provide detailed holdings information. 10.1 million links to the German Integrated Authority File (Gemeinsame Normdatei – GND) – an authority file for persons, corporate bodies, conferences and events, geographic information, topics and works[1] – underline not only the

[1] For further information regarding the GND see http://www.dnb.de/EN/Standardisierung/GND/gnd_node.html

high degree of standardisation of the ZDB but also the high quality of its data. The latter is also determined by the comprehensiveness of the provided information and its machine readability. This quality allows for various visual representations of the information, making it more explicit, enabling queries and explorative search processes that are supported in intuitive, interactive ways. For searches regarding newspapers and corresponding digitisation projects, the following title-associated information is of main interest:

- Title relations
  Interconnections between titles are exposed in form of links between title records in the ZDB. This data can be utilised for displaying in a graph clusters of titles that are connected directly or indirectly. The following table shows the main title relations in the ZDB and their frequency of occurrence.[2]

Title relations	Frequency of occurrence (September 2015)
Chronological relations (predecessors, successors)	850,000
Relations “has supplement”	83,000
Relations “is supplement of”	74,000
Parallel editions	310,000

- Time-related information
  Publishing dates and dates of volumes and issues that institutions have in stock are stored as machine-readable data. Modern search interfaces may use these to support queries for exact years as well as searches based on timelines.

- Geo-coordinates of places
  Some geographic GND data contain coordinates from the Geonames database.[3] These can be utilised for representing distribution places of newspapers in a map. In the ZDB, about 20,000 places of distribution are linked to GND records. An analysis of the ZDB and GND data has shown two things:
  1. Not all GND records to which corresponding places of distribution in the ZDB are linked contain geo-coordinates.
  2. Relatively many titles are connected with relatively few places of distribution (see the following diagram).

[2] The table does not include interrelations that were introduced with the cataloguing standard *Resource Description and Access (RDA)*. With RDA, an enhanced set of relationships is available that allows for title relations being represented more precisely than before. For general information about RDA see for example: Mering, Margaret (ed.): The RDA workbook : learning the basics of Resource Description and Access. Santa Barbara, Calif. ; Denver, Colo. ; Oxford : Libraries Unlimited, 2014.
[3] http://www.geonames.org/

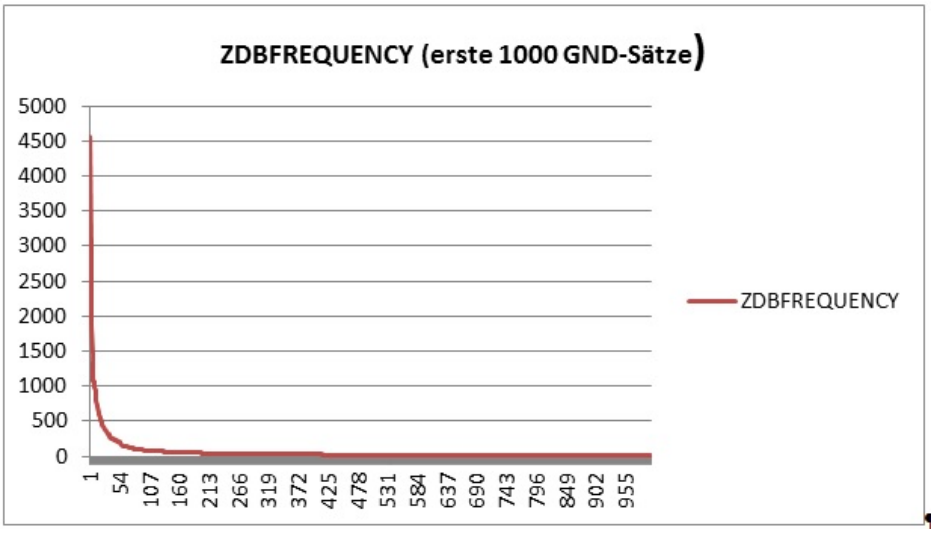

Diagram: Correlation between ZDB title records and places of distribution.
X-axis: number of ZDB titles, Y-axis: connected GND places

In the last year, the DNB has added geo-coordinates to 200 records of the GND that are linked to the largest number of titles. As a result, the GND now contains sufficient information that can be used for the display of places of distribution on an interactive map.

The current ZDB OPAC [4] follows the traditional OPAC concept, focusing on single bibliographic records, not offering any visualisations for supporting searches or exploration at all. In the last years new search concepts have evolved that take contexts of information much more into account and understand graphical representations to be suitable means for supporting specific research questions. Presenting direct and indirect relations between titles as graphical networks is an old requirement that has been formulated years ago[5] and that finally could be realised last year within the frame of a project that aimed at creating an entirely new ZDB catalogue.

2. The new ZDB catalogue

2.1 Project context

The first beta version of the new ZDB catalogue was developed within a project that aimed at enabling the ZDB to serve as the main steering and research instrument for newspaper digitisation projects and for newspapers in Germany. The project was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft – DFG) and ran from February 2014 to June 2015. It was part of a joint project investigating aspects of the digitisation of historical newspapers that ran from May 2013 to December 2015. Six libraries were involved in the joint project: Bavarian State Library, German National Library, State and University Library Bremen, Berlin State Library, Saxon State and University Library Dresden, University and State Library Halle. In the joint project, different organisational and technical digitisation methods were tested and evaluated. Additionally, common presentation and search environments were improved: On the one hand the DFG viewer – a tool for presenting

[4] http://dispatch.opac.d-nb.de/LNG=EN/DB=1.1/
[5] See for example Bernd Augustin and Johann Rolschewski: Zeitungen in der ZDB. In: Zeitschrift für Bibliothekswesen und Bibliographie 57 (2010), Nr. 3/4, pp. 152 – 155. P. 153f.

of and navigating in digital content – was adjusted to the specific requirements of newspapers. On the other hand, a beta version of the new ZDB catalogue was created that provides intuitive search interfaces and useful interactive graphical representations of the data. Results of the joint project and insights gained were enriched with a set of recommendations for a national funding scheme for newspaper digitisation and submitted to the DFG.[6]

The German National Library and the Berlin State Library assumed responsibility for creating a new ZDB catalogue that could function as the major data infrastructure of such national funding scheme. The development of a new (mostly barrier-free) front end with a completely overhauled new design was done by the]init[AG.[7] With the publication of a first beta version of the new ZDB catalogue, the project was officially completed. However, work on the new ZDB catalogue is still in progress as it is planned to substitute the current ZDB catalogue within the next year.

2.2 Features of the new ZDB catalogue

The new ZDB catalogue does not only offer a newly designed search interface – including a responsive design adjusted to tablets and smartphones – but also several new features to support title queries:

- An autosuggest function assists users in formulating search queries.
- Applied search terms remain visible in the search field on almost every page of the catalogue. Queries can thus be easily reformulated at any time if obtained results are not satisfactory.
- Faceted search options, checkboxes and a graphical timeline help users refining search results. Selected attributes are visible in a breadcrumb trail displayed on the results page and can be easily deselected.
- Direct and indirect predecessors and successors of a title can be viewed and searched for in a "title history". The feature "title relations" offers information about further interrelations between different titles. Both visualisations include a zoom function and may be expanded to a full screen presentation. Graphic files also may be downloaded as png or printed. Further characteristics will be described in detail later on.

Additionally, the new catalogue supports searching for titles in nonlatin original scripts. Results may be optionally viewed in nonlatin or in latin script.

Like the current ZDB OPAC, the new catalogue offers sorting options for result lists, a search history and a watchlist. Single title records can be downloaded in different formats and may be ordered via the menu "option" in the single title view. The new catalogue also presents holding and license information besides title data, particularly the holding data is now much more clearly arranged than in the old catalogue. The new ZDB catalogue offers several new features to support queries for holdings:

- Holdings can be sorted according to place of library, international library code (ISIL[8]) and interlibrary loan region.
- Holdings can be filtered via facets.

[6] For more information about the joint digitisation project see Thomas Bürger: Newspaper Digitization in Germany – Results and Perspectives of the German Research Foundation Pilot Scheme 2013-2015. In: IFLA International News Media Conference 2016.

[7] https://www.init.de/en

[8] "International standard identifier for libraries and related organizations". For more details see ISO 15511 (http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=57332)

- In a feature called “holdings overview” library holdings of a specific title can be compared. This feature will be described in detail later on.
- Libraries with holdings of a required title may be searched for via a map. Geo-coordinates stored in the ISIL agency[9] are used for this purpose. The map is offered in combination with a timeline that enables users to restrict the holdings view to a specific time period. A grey shadow in the timeline indicates volumes that are available in the institutions visible on the map. The following figure gives an impression of this feature.

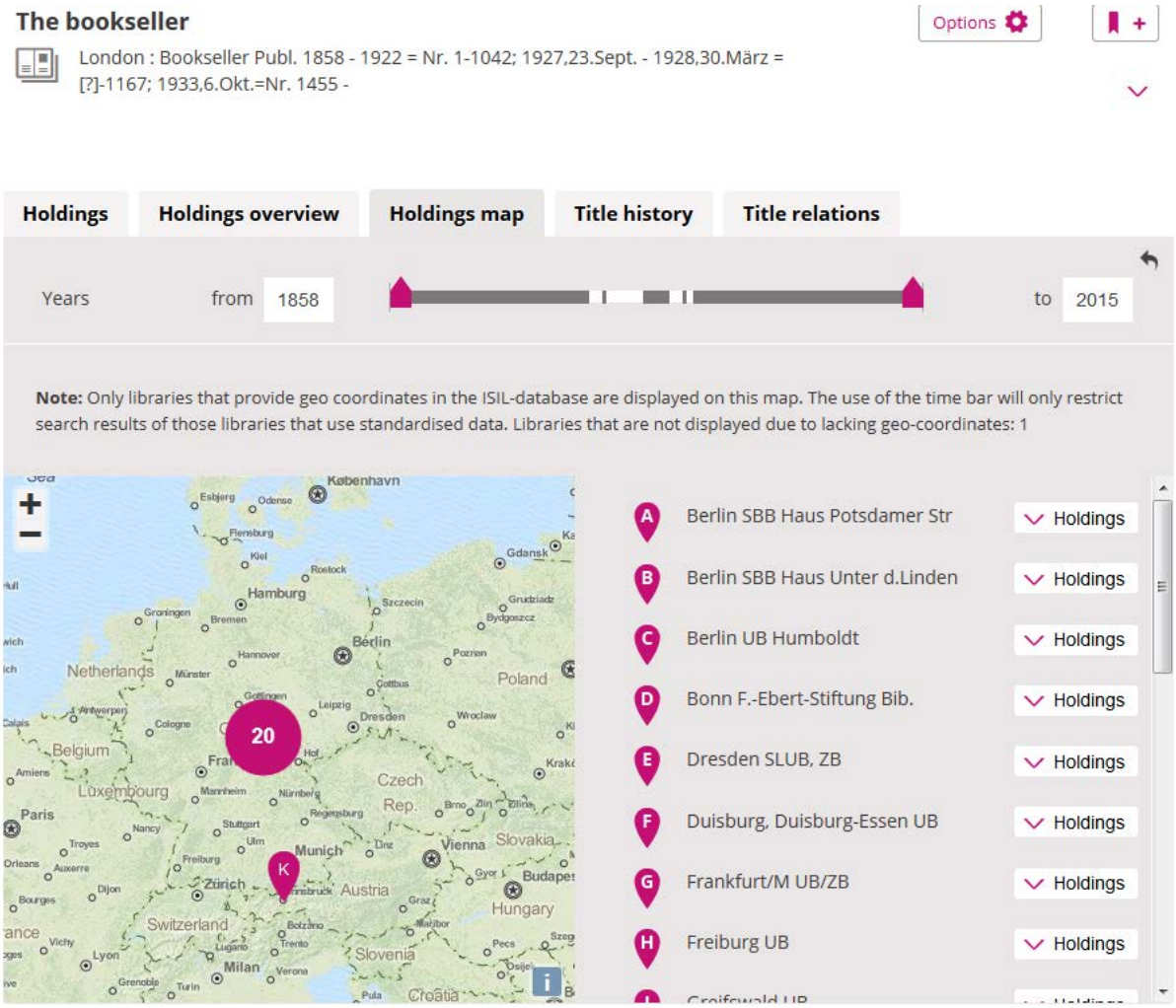

Fig. 1: Holdings map for “The bookseller”

- Relevant information for a specific library can be found on a specific location map (see figure 2)

[9] http://sigel.staatsbibliothek-berlin.de/en/startseite/

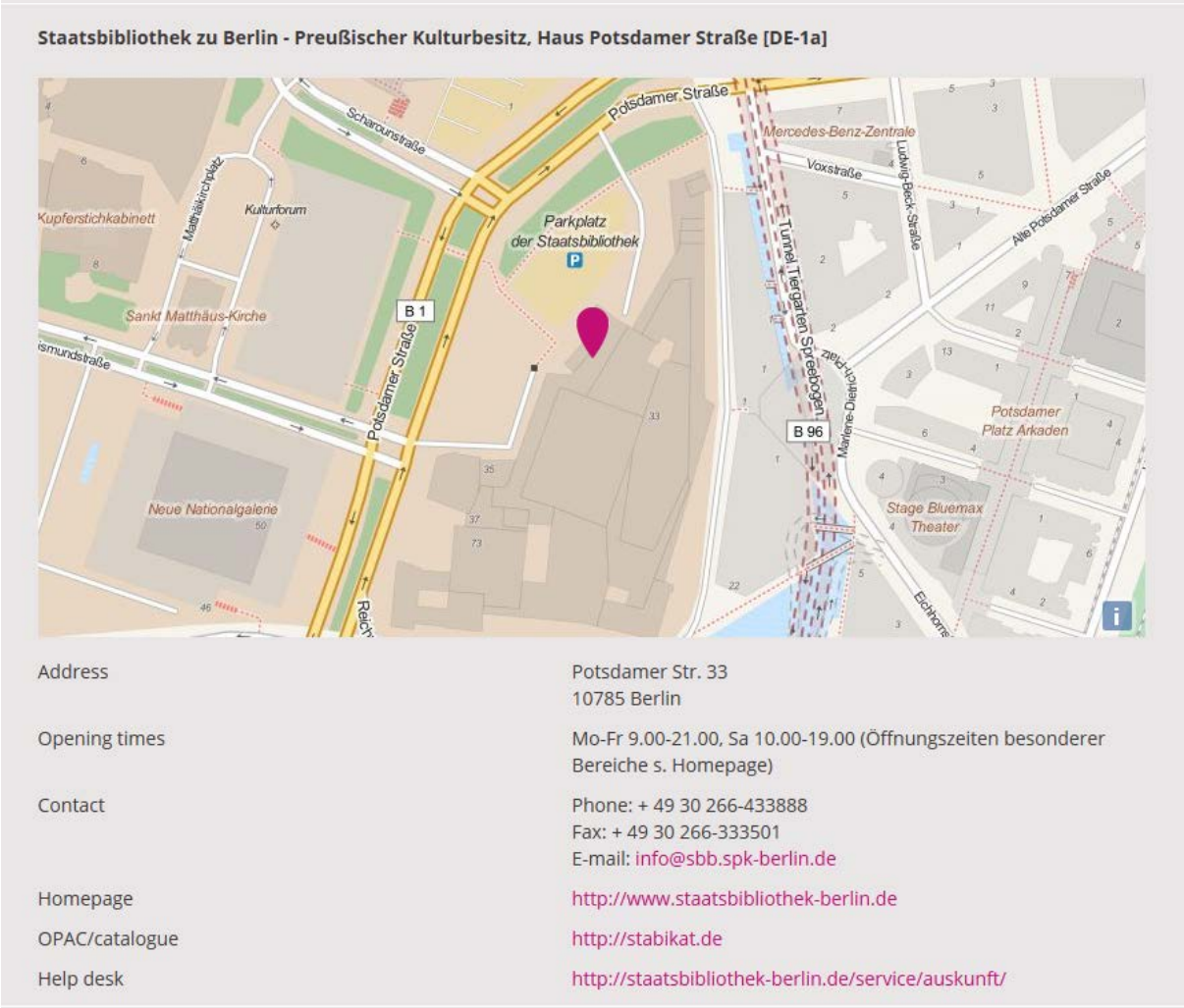

Fig. 2: Location map for the Berlin State Library (Staatsbibliothek zu Berlin)

3. Visual representations in the new ZDB catalogue

In this section, the following features are described in detail:

- Holdings overview
- Title history
- Title relations.

All three features are offered via tabs in the single view of a title, provided the corresponding data is available.

3.1 Holdings overview

In large-scale co-operative digitisation projects with several participating institutions, the identification of specific holdings in libraries can be a challenging and time-consuming task. In the past, the ZDB team at the Berlin State Library selected and evaluated the necessary information on request by the partner libraries. With the new ZDB catalogue, digitising institutions can easily inform themselves about the availability of print originals via the graphical holdings overview.

In the holdings overview the volumes of a title that an institution has in stock are visualised in a timeline – provided the institutions have catalogued the corresponding information in the ZDB as machine-interpretable data. With the timelines it is possible to see and understand at a glance which library has the most volumes of a title in stock or, equally interesting, owns particularly rare volumes. The holdings overview consists of three elements (see also fig. 3):

- General holdings filters
- Overall timeline for holdings overview
- Graphical representation of the volumes of the title that institutions have in stock.

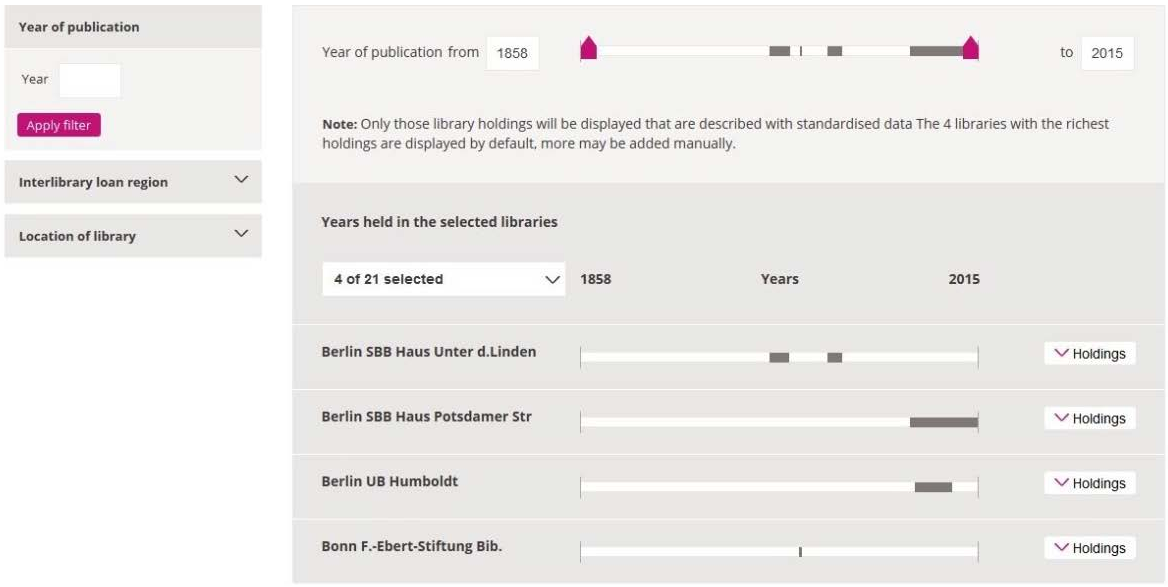

Fig. 3: Holdings overview of “The bookseller” (http://beta.zdb-opac.de/zdb/title.xhtml?idn=012416673)

The holding filters are displayed on the left hand side and are available in all holding views (Holdings, Holdings overview, Holdings map). They can be used to restrict the shown holdings with regard to a specific volume, one or more interlibrary loan regions and one or more locations of institutions. On the right hand side the actual holdings overview is displayed. On top, a general timeline informs users in visual form about the combined volumes the selected institutions displayed below have in stock. Input fields for years as well as sliders allow for filtering to specific time spans. Volumes that individual institutions have in stock are visualised in individual timelines next to the name of the institution. Each institution appears in a separate line. In this way it can easily be evaluated which library has desired volumes in stock. A button “holdings” next to the individual timelines expands to detailed holdings information in written form (see figure 4). By default the holdings of four institutions are displayed in the overview; others may be easily selected or more added.

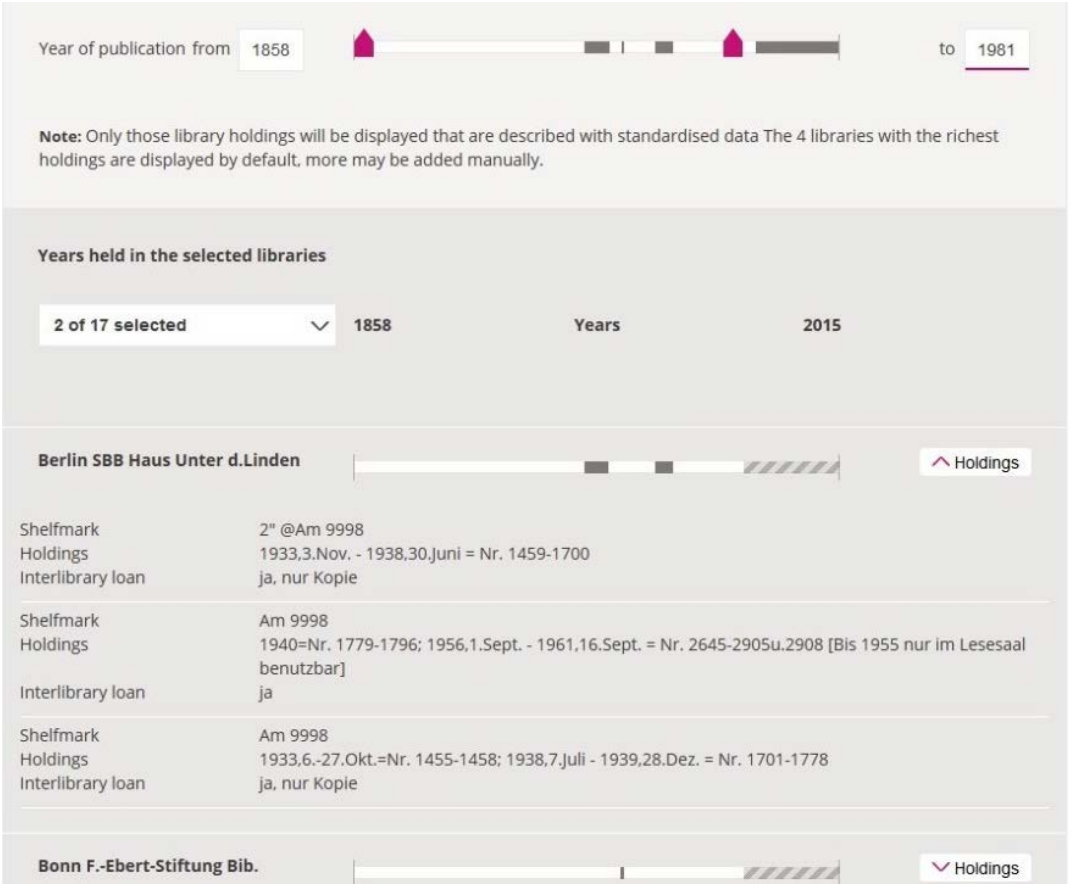

Fig. 4: Holdings overview of “The bookseller” with detailed holdings information for the Berlin State Library (http://beta.zdb-opac.de/zdb/title.xhtml?idn=012416673)

3.2 Title history

The title history provides an overview of the development of a newspaper over time by displaying its direct and indirect predecessors and successors including the respective years of publication in chronological order. By default the entire span of a title history is displayed. A timeline allows for the selection of a specific time span of a given title history. A grey shape in the timeline indicates the part of the history that is visible on the screen, a helpful orientation especially in case of long histories. If users place the mouse over a displayed title, its directly linked predecessors and/or successors are highlighted. A click on a title will link to its single view. The title that constitutes the initial user access to the title history is highlighted in magenta and is, of course, not provided with a link. Figure 5 gives an example of a title history.

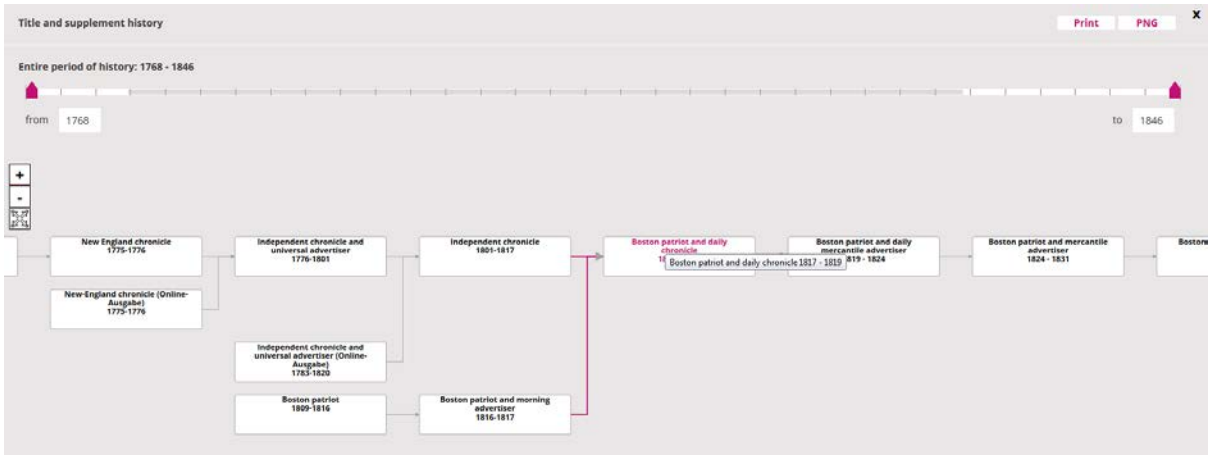

Fig. 5: Extract from the title history of “Boston patriot and daily chronicle” (http://beta.zdb-opac.de/zdb/title.xhtml?idn=014500930)

Newspapers with supplements are equipped with a plus-symbol in the title history that expands to a visual display of the publishing periods of the corresponding supplements, provided the corresponding years are available in machine-readable form. Each supplement is presented in a separate line and provided with a link to the single view of the corresponding supplement. At the top of the supplement history a timeline informs about the publication years of the newspaper title. Users can easily see whether a title was published continuously or not by looking at the grey timeline in which white spaces indicate gaps in the publishing history. Fig. 6 gives an example for a supplement history.

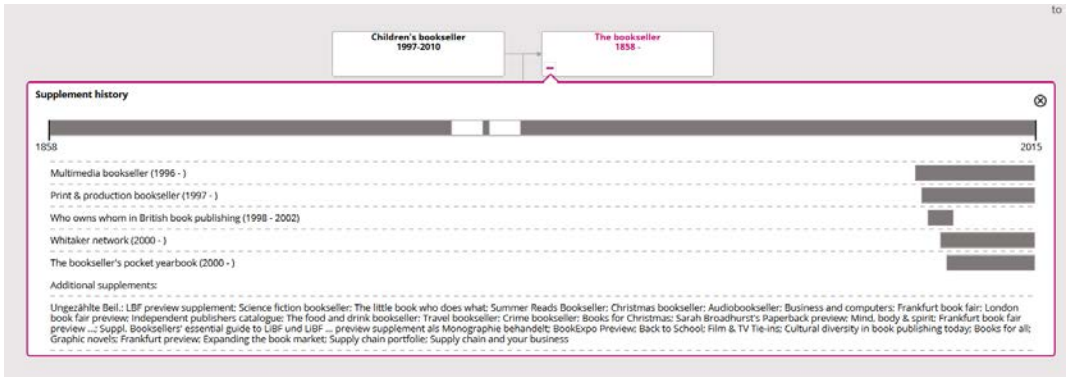

Fig. 6: Supplement history of “The bookseller” (http://beta.zdb-opac.de/zdb/title.xhtml?idn=012416673)

3.3 Title relations

Whereas the title history focuses on predecessors and successors of a title (and included supplements), the feature “title relations” displays the entire network of titles being connected to an initial title, provided they are adequately documented in the ZDB. Relations are represented as unidirectional links starting from the initial title that is displayed in magenta. The relations are described in unambiguous terms that reflect their directionality:

- has predecessor
- has successor
- has supplement
- is supplement of

- has parallel edition

Different relations are marked with different colours. If several titles are linked via the same relationship, they are presented as a cluster. All titles – except the title that constitutes the initial user access to the title relations – link to the respective single views. Figure 7 gives an example for a relational network being available in the title relations view.

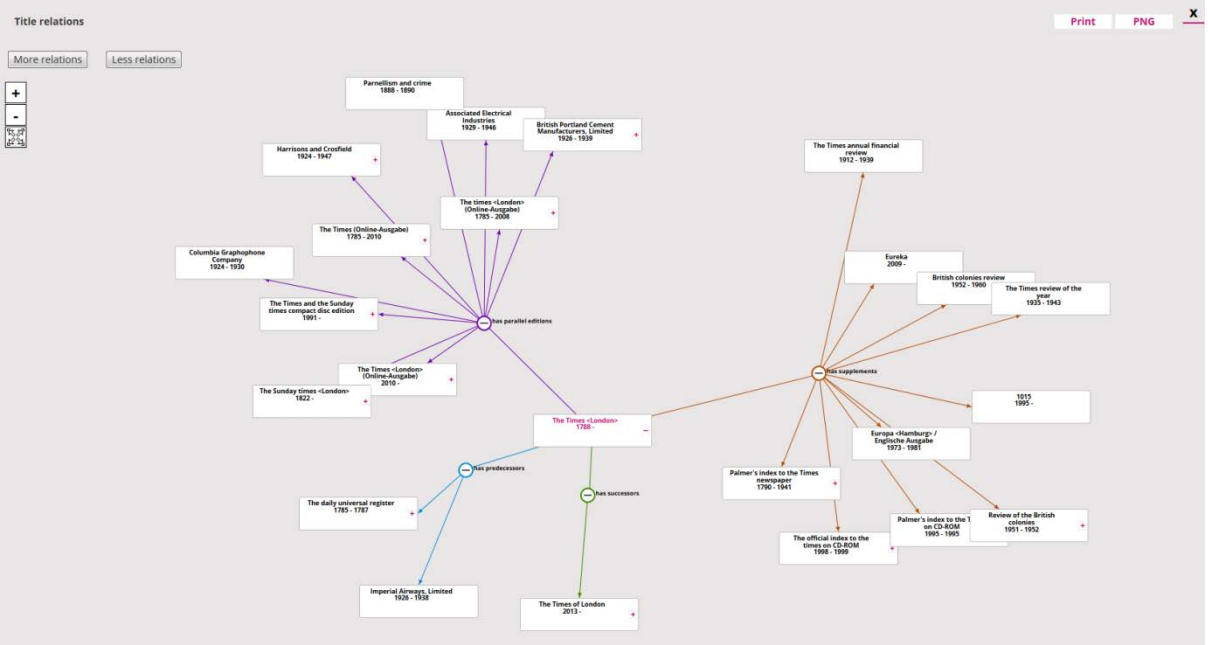

Fig. 7: Title relations for “The Times” (http://beta.zdb-opac.de/zdb/title.xhtml?idn=010716106)

Since title relations can form amazingly complex networks, three mechanisms for expanding and collapsing the displayed graph are provided:

- The graph can be gradually expanded or collapsed by clicking on the buttons “More relations” and “Less relations”.
- Plus- and minus-symbols, if available, expand and collapse the nodes of the network.
- The graph may also be partly expanded or collapsed by addressing specific relationships, e.g. show all parallel editions, show all supplements etc.

Users may apply these mechanisms in combination, which allows them to create graphs of a complexity adjusted to their needs. Additionally, all nodes of the graph may be freely arranged on the screen for an optimised user experience. By default, the direct relational context of a title is presented.

4. Summary and Outlook

Doubtlessly, the interactive visualisations constitute the most prominent features of the new ZDB catalogue. They do not only support search interactions including explorative approaches in a modern way but also help in utilising the information so far “hidden” in single data records, for making larger contexts explicit and visible. Yet, these are not the only new features that simplify searches. A checkbox “planned digitisation” on the enhanced search site allows for the easy identification of titles intended to be digitised. Additionally, result lists may be reduced by a variety of criteria: publishing type (e.g. newspaper), publishing frequency (e.g. daily), form of resource (print, electronic, digital), country and language of publication – all available as search facets.

One of the next versions of the new ZDB catalogue will include a new facet “place of distribution” combined with the option to look at and search for distribution places of newspapers in a map. For this feature, geo-coordinates stored in the GND will be used. Figure 8 gives a preview of this feature.

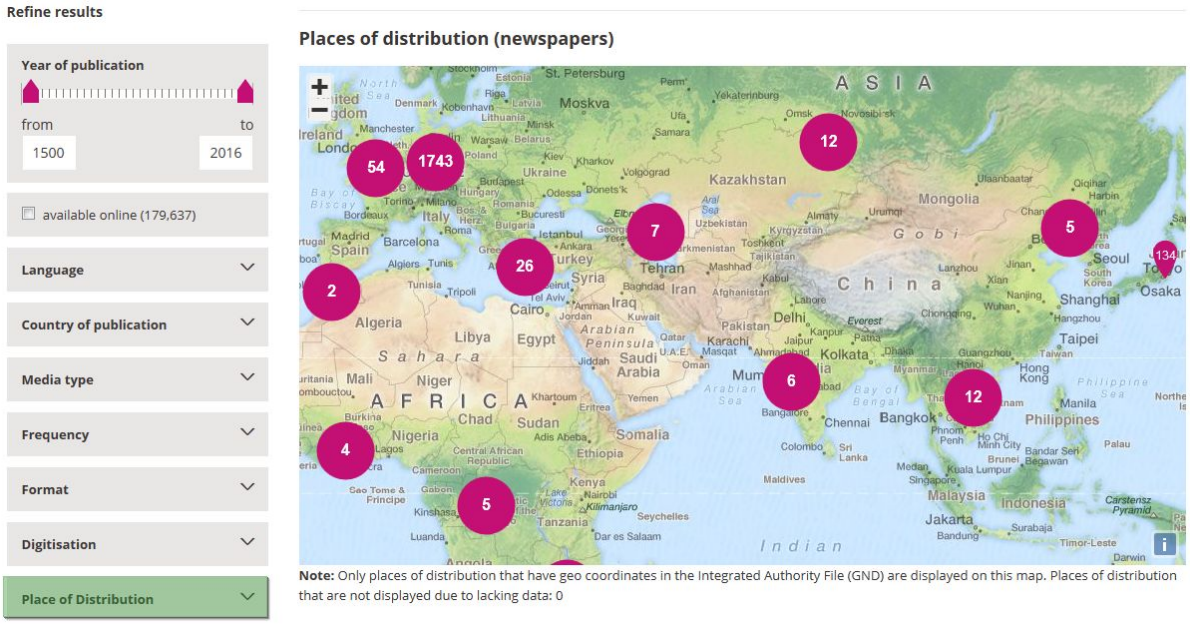

Fig. 8: Upcoming visualisation: facet & map “places of distribution”

In this paper existing features of the new ZDB catalogue were mentioned, further ones are under development, and the new catalogue is constantly being improved. The beta version is intended to be replaced by a productive version that will also replace the old ZDB OPAC in 2017. Since the public beta version already accesses all ZDB data, obtained search results can be regarded as reliable and indicative.