

Collection mandate of the German National Library

"The extensive collections of text, image or sound recording-based publications issued since 1913 in Germany or in German, constitute a comprehensive, objective, unbiased and permanently preserved representation of a major part of the cultural heritage of Germany based on the library's legal mandate." This guiding principle heads the "Collect and Preserve" section in the Strategic Compass of the German National Library¹, although the validity of this statement extends beyond the ten-year horizon of the strategy. From the outset the German National Library has repeatedly had to update its collection mandate in response to economic, scientific and social developments. In the face of the digital revolution, the need to describe and plan the execution of the mandate is more obvious than ever.

The legal mandate of the German National Library is to collect, index and preserve everything that has been published in Germany, about Germany and in German since 1913.² This includes all editions of sheet music and music resources published in Germany. The legal deposit regulation for notated music and music records has only been in force since 1973, however the collection stretches back retrospectively to the beginning of the music recording age at the end of the 19th century. The German National Library is thus Germany's cultural memory for the written word and for music of the recent past and the present. But what functions should a national cultural memory fulfil in the era of the World Wide Web (WWW), digitisation and virtualisation?

This question is being asked as part of the wider debate on the consequences of the digital transition which is still in full swing. Those responsible in the German National Library agree that it is only possible to answer this question as the situation develops and by engaging in an ongoing dialogue. The German National Library has thus conducted four loosely connected workshops³ aimed at discussing the issues with representatives from other institutions and members of its advisory boards, and at encouraging further development of the collection guidelines⁴.

The following reflects the current status of activities and the position of the German National Library, while taking into account the results of the workshops.

¹ German National Library 2025 - Strategic Compass. Leipzig, Frankfurt am Main, 2016. ISBN 978-3-941113-45-9, <https://d-nb.info/1112299556/34>

² See appendix: "Auszug aus dem Gesetz über die Deutsche Nationalbibliothek" (= Excerpt from the Law regarding the German National Library (DNBG))

³ "Dynamisches Bewahren" (Preserving the Dynamic) (2013), "Musik im Netz" (Music on the Internet) (2014) and "Digitales Gedächtnis" (Digital Memory) (April 2016) which was supplemented by an additional session with the regional legal deposit libraries (August 2016).

⁴ Such collection guidelines are also used to document and trace decisions on what is included in or excluded from the collection over time.

First steps: Focus on commercial and scientific publications

The amendment Act of 2006 tasked the German National Library with responding more actively to digital developments: since then, the collection of media works published online has been an explicit part of its tasks. The law as well as the legal deposit regulation deliberately do not address the question of how this mandate is interpreted in practice⁵. However, all content disseminated on the publicly accessible Internet basically falls within the collection mandate. The German National Library therefore began to extend its collection guidelines at an early stage to include online publications and in 2009 issued initial details of which digital resources were to be collected and which not. From the outset a distinction was made between digital publications collected as individual objects and typically published by publishers, music labels, universities or other research institutions on the one hand, and digitally-born content created with and in the World Wide Web on the other, which are gathered as whole collections, i.e. within their overall context on the site. The German National Library proceeded highly pragmatically, starting with publications from publishers and institutions with which it already had close links. In addition, it was and remains necessary to continue fleshing out the basic, abstract concept of the German National Library with regard to the collection of online publications and to make it more practicable.⁶

Even before the law was amended, the German National Library was able to gain initial experience with this form of online publication on the basis of voluntary deposits by publishers and universities. Since then it has expanded its expertise considerably and continuously in this field. Yet the number of publications, especially non-commercial and non-institutional, continues to grow steadily.

Scientific periodicals, non-commercial scientific publications as well as "self-published" works are affected in particular by the rapid increase in online publications. These publications, which cover all conceivable subject areas and areas of knowledge at different quality levels, characterise an ever-expanding field of German culture. It is of paramount importance to ensure that both these and the physical media works remain permanently available. Otherwise, gaps will arise that make a comprehensive description of cultural, scientific and social developments and analysis of their causes and effects difficult or impossible in the future.

On 1 October 2016, the online publication holdings of the German National Library included:

1,305,261 monographic online publications
including

- | 649,954 e-books (mainly commercial publications)
- | 450,221 self-published titles (BoD)
- | 188,695 online dissertations and other university publications
- | 11,569 audio files (audiobooks)
- | 4,822 digital notated music works

3,931 e-journals (current and completed titles) with 603,013 issues or articles
1,304 e-papers (titles) including 1,280,758 individual issues
1,920 harvested websites (regular or event-related collection)⁷

⁵ See appendix: "Auszug aus dem Gesetz über die Deutsche Nationalbibliothek (DNBG) (=Excerpt from the Law regarding the German National Library (DNBG))

⁶ See "Sammelrichtlinien, Stand 1. Mai 2014 (= Collection Guidelines, as of 1 May 2014), pages 49-61 and 69-71 at <http://d-nb.info/1051940788/34>.

⁷ For current holding figures and the development of inventoried accessions since 2014 see: Annual Report 2016 of the German National Library at: <http://d-nb.info/1135316724/34>, p. 46-49.

This statistic clearly highlights the phased approach of the German National Library:

in contrast to the approach of other national libraries, e.g., in France, where the collection of online publications began with web harvesting, the focus of the German National Library in the early years was on the digital output of publishers and universities for which a counterpart exists in physical form or is readily conceivable – i.e. on institutions and content with which the German National Library has enjoyed a long history. Here the objective of the German National Library is to establish uniform, efficient mass delivery procedures for electronic production as a means of achieving a complete collection. The corresponding procedures were initially (and are still being) developed and tested with selected partners, and then transferred step by step to other suppliers and/or other types of media. The first stage was to develop procedures for monographs (e-books) and university publications (e.g. online dissertations and other university publications).

Building on this, the procedures for e-papers (PDF editions of daily and Sunday newspapers) and other serial resources (e-journals) were then expanded.

Technical preparations for web harvesting were only started in 2010; the first crawls were carried out in 2012.

With the exception of e-papers (nearly all of which are now collected) and e-books (around 80 per cent of which are currently being collected), collecting activities for the other categories of online publications mentioned above (only specific forms of which were initially collected: commercial and scientific publications) must now be significantly expanded. A comparison with the German Union Catalogue of Serials showed that the German National Library's collection rate of e-journals at the end of 2015 was approximately 12 per cent. In the area of digitally distributed notated music and music recordings, essentially only music resources deposited by publishers via the interface on the German National Library website have been collected up to now.

Development of a WWW collection strategy

The boundaries between different media forms, communication and publication, and the public and private sectors are more blurred on the WWW than is the case for physical media works. The German National Library is therefore having to work harder than ever before to fulfil its mandate here. There are enormous technical, conceptual, organisational, financial and legal challenges to overcome. New methods, infrastructures and processes must be developed, binding regulations formulated and appropriately qualified employees taken on to cope with the wide range of digital products. Established procedures can only be used to a limited extent, meaning that these tasks can only be accomplished gradually and by defining clear areas of focus.

After the legal mandate was expanded to include online publications, the German National Library therefore initially excluded certain forms of publication from its concrete collection concepts and their implementation. These include specialist or political Internet forums, blogs, communication methods such as social media or new forms of presentation of commercial or non-commercial bodies on the WWW, such as digital reworkings and presentations of specific (physical) collections on websites (e.g. preparation and presentation of SED files by the Federal Archives in the SAPMO project), in virtual exhibitions (e.g. the online database of the collections of the Staatliche Museen zu Berlin, SMB-digital) or similar forms. Instead, it has limited itself to gaining initial experience with the current possibilities of web archiving by taking small steps in collaboration with a service provider. These activities are important in order to gradually accrue the necessary technical expertise and to be able to make a realistic assessment of existing content, the work and costs involved, any prerequisites and the feasibility of particular activities. This approach consciously accepts that much potentially collectable content will disappear in the meantime. However, a recent study revealed that the content behind quoted URLs has changed since the time of citation,

or is no longer accessible, in seven out of ten scientific, technological, and medical articles citing online sources.⁸

The German National Library searched for solutions in the aforementioned workshops and discussions with experts. It formulated its strategic guidelines for dealing with this collection area on the basis of these exchanges and its own practical experience, (see also "German National Library 2025").

In short, the German National Library intends to respond to the inherently ephemeral nature of the Web by adopting two different parallel collection strategies in order to fulfil its heritage mandate for future generations: these involve taking two types of "snapshot" – "**selective harvesting**" and "**de-crawls**" – in an effective, economical and evaluable way. In both cases, openly accessible and free sites are "collected" from the WWW.

- | Content selected by topic or institution is collected systematically at regular intervals. This selective web harvesting allows the crawl frequency to be set for each individual site. The German National Library originally adopted a quarterly frequency but moved to a semi-annual frequency based on the experience of the first two years. This was because it discovered that most of the sites it was collecting (government agencies, associations, etc.) only changed relatively infrequently. A higher frequency is set for websites that change more often.

More than 1,200 websites have been regularly collected since 2012⁹. Event-related crawls are also performed¹⁰. There were a total of 9,896 individual snapshots stored in the web archive on 1 October 2016¹¹.

The number of selectively harvested websites will need to increase significantly over the next few years in order to ensure that sufficient cultural, scientific, social, political, religious or personal testimony of the present day is kept for posterity. The present target is an annual increase of 25 per cent on the current number. In 2017, the German National Library published its Strategic Priorities for the period 2017-2020. Web harvesting represents an important task area. It includes reviewing, revising and expanding existing criteria for selective harvesting. Further aspects include analysis and any necessary adaptation of infrastructure and resources. Communication measures are also important to ensure transparency in the collection policy.

- | As a cross section, "snapshots" of the Web are also collected based on purely formal criteria. This provides an overall view of the "German" Web and fills at least some of the gaps that arise naturally through selective web harvesting. They also allow for random finds which are not included in the results of selective harvesting searches. The frequency of these ".de crawls" is dependent on financial and technical (storage) framework conditions, yet must yield a coherent picture of the development of the "German" Web, rendering it usable for historical and scientific analysis in the future. The German National Library has only had one such ".de-crawl" carried out so far, in 2014. Approximately 6 million websites including a total of approx. 2.5 billion files (120 TB) were collected. These

⁸ Quoted from: Tobias Beinert, Astrid Schoger: Vernachlässigte Pflicht oder Sammlung aus Leidenschaft? Zum Stand der Webarchivierung in deutschen Bibliotheken. In: ZfBB 62. 2015, 3/4. - p. 172-183

⁹ Websites of federal authorities and institutions, interest groups, cultural institutions, sports associations, religious communities, political parties, party-related organisations and politicians, scientific societies, research institutions, social security institutions and selected specialist topics (for example long-term digital preservation) are collected regularly.

¹⁰ Event-related crawls were made e.g. for the 100th anniversary of World War I, the 200th birthday of Richard Wagner, the flood in 2013, 25 years after the fall of the Berlin Wall, 50 years of diplomatic relations between Germany and Israel, the 2013 federal election, FIFA World Cup 2014, 25 years of German unity, the refugee crisis in Germany from 2015 and the Jan Böhmermann case.

¹¹ There are two cost factors: the number of individual crawls and the amount of data that is hosted and made available. Currently, the German National Library has agreed a fixed price of EUR 108,900 (net) per year with the service provider, including 4,000 crawls per year and 14 TB of storage volume.

are now stored by a service provider. Regular iterations of such ".de-crawls" are regarded as meaningful for the future.¹²

The German National Library is convinced that both approaches (selective harvesting as well as domain crawls at defined time intervals) are needed in order to preserve the cultural heritage through comprehensive and neutral sampling – precisely because the borders are blurring on the Web, which often makes it impossible to distinguish between private and public information without great effort.¹³

It is important to enter into (or to expand existing) partnerships with other institutions here. Discussions are therefore being held at the international level with the Internet Archive into possible synergies in the selection, collection and presentation of German websites. On a national level, the regional legal deposit libraries represent obvious partners with which sensible and cost-effective division of labour is conceivable, provided legal issues can be resolved to permit collaboration which is not burdened by complex trilateral contractual agreements. Currently, the German National Library organises the selection and prioritisation of the selective collection of websites in open and public dialogue with its funding institution, with politicians, academics, other cultural institutions and relevant experts: it requires continual support in resolving the issues which constantly arise and therefore intends to proceed with the series of workshops it has already begun.

Digital publications (individual objects) from publishers, music producers, public authorities and large institutions are announced and disseminated via channels that the German National Library has used and evaluated over a period of time. However, there is a large number of works on the Web about which these channels provide no information. The goal of completeness should be abandoned for those parts of the web that may be relevant for digital collection; representativeness should be sought instead. This selection and representativeness should be defined in collaboration e.g. with the Länder libraries in Germany or with other national libraries. Representativeness means enabling future generations to reconstruct a picture of a particular time through selective web harvesting as well as ".de crawls", i.e. by undertaking regular, nationwide crawls and consciously defining collections that are considered important and typical of the present day. In the past it was often left to coincidence to decide whether a piece of cultural testimony was preserved and passed down, whereas now this process is to be consciously managed with the help of transparent selection criteria so that future generations have the necessary sources and consciously permitted gaps can be documented in an accountable manner. Which content of the WWW is worthy of public interest and should thus be archived must also be examined and discussed. The German National Library has a responsibility to collect its holdings in a way that provides a comprehensive and unbiased overview. The idea of a selective but nevertheless representative collection of web content organised in concert with other institutions and individuals must then be reconciled with the role of the German National Library as an institution that

¹² The service provider charged the Library a total of EUR 112,500 (net) in the first year (2014) for the 120 TB ".de crawl", with annual hosting costs of EUR 50,000 (net) for each subsequent year. A follow-up crawl, it is estimated, would harvest approximately 200 TB of data, resulting in a corresponding increase in costs. Added to this are the current annual costs for hosting.

The Bibliothèque nationale de France (BnF), for example, is pursuing a similar strategy. At present, it conducts one domain crawl a year; this resulted in the collection of approximately 4.4 million websites in 2015 alone. Furthermore, thematic crawls of around 20,000 websites are conducted several times a year. In total, the BnF's internet archives now contain 26 billion files, some of which date back to 1996.

Source: http://www.bnf.fr/en/collections_and_services/book_press_media/a.internet_archives.html (last visited on 18.10.2016)

¹³ In mid-October 2016 the holdings from regular and event-related crawls comprised the following: 1,306 crawls of federal authorities and institutions, 1,749 of interest groups, 1,039 of cultural institutions, 851 of sports associations, 681 of religious communities, 859 of party-affiliated organisations and politicians, 458 of scientific societies, 493 of research institutions, 289 of social security institutions, 19 on the subject of "Digital preservation" and 632 in the category "Miscellaneous".

preserves relevant German cultural heritage authentically and reliably without any judgement of the content.

The current collection guidelines of the German National Library already distinguish between online publications which are to be collected in full (individual objects) on the one hand, and Web content selections (collections) on the other. These are continually being refined.

The current collection policy for web harvesting is being reviewed in 2017 and, if necessary, new or modified selection criteria will be defined regarding which websites the relevant software programs (crawlers) collect, including specification of the time frame and budget (see below). These specifications must be reviewed at regular intervals and, if necessary, adapted to reflect new developments. This also includes the question of whether and to what extent formats not previously collected (such as blogs) should be included. The reviews can also ascertain whether the collections contain individual objects that should be addressed and catalogued as such.

With regard to web content selection, the question also arises of whether the existing roles of museum, archive (text, music, film archives) and library can be transferred to the WWW. This fundamentally different approach to the collection mandate became clear at the "Digital Memory" workshop on 12 April 2016. While legal deposit libraries consider the collection of Web content part of their legal mandate, and enjoy privileged rights in the collection of digital works over other, smaller and sometimes private, institutions, e.g. with regard to copyright law, representatives of archives and museums point out that they only archive websites if there is a connection to other material which falls within their collection mandate. As a general rule, they believe that responsibility for the collection of websites lies with the German National Library and the regional legal deposit libraries. Collection of their own digital publications on the Internet (virtual exhibitions, collection presentations) is oriented towards individual objects, and not based on the form of presentation in the Internet. A glaring difference has emerged here in the understanding of what constitutes a digital publication; the German National Library needs to set out and communicate its understanding of "publication".¹⁴

The fact that much digital content is highly ephemeral is an important argument in favour of intensifying the web harvesting activities of the German National Library. In a recent comparison of the holdings of harvested pages from the ".de" domain of the Internet Archive with those of the only comprehensive crawl of the German National Library so far, it was found that the Internet Archive has saved only about 38 per cent of the pages collected by the German National Library.¹⁵ In this respect, there is increased pressure for the German National Library to act with regard to the collection of web pages, since the often transitory content has not been submitted to any other institution. This pressure also extends to agreements on partnerships with other institutions, on data transfers or the division of responsibilities.

Possible cooperation between the German National Library and the regional legal deposit libraries was discussed at a workshop in August 2016. Two models were suggested by the regional depository libraries: either the regional libraries select the content which is then harvested and archived by the German National Library, or both the content selection and archiving are distributed based on agreements. Guaranteed access is a prerequisite for both models, or for a combination of both models – at least for all libraries that have legal deposit status.

At its meeting on 5 December 2016, the Administrative Council of the German National Library also discussed the topic of webpage collection by the German National Library and unanimously adopted the following resolution: "*The Administrative Council supports a selective and cooperative approach*

¹⁴ see Appendix: "Definition des Begriffs „Digitale Publikation“ und aktuelle Verwendung der Terminologie in der Deutschen Nationalbibliothek" (=Definition of the term "digital publication" and the current use of terminology by the German National Library)

¹⁵ The German National Library also plans to clarify what percentage of the pages collected by the Internet Archive also fall within its own collection mandate but has not been collected.

to the collection of web pages in the collection concept put forward by the German National Library. It requests that the German National Library promote the discussion of possible selection criteria with suitable partner institutions.

Furthermore, it agrees to perform qualitatively low level crawls of the ".de" domain at specified intervals. This practice should be reviewed in terms of data protection and copyright."¹⁶

Next steps in the collection of online publications

In parallel to discussing the above questions regarding the collection of web-specific online publications, the German National Library must continue to work on building its collection of individual objects. It maintains its goal of completeness for commercial, scientific and music industry publications. This also includes scientific open-access publications. The sometimes low degree of coverage (see above) reveals that considerable work is still needed here.

The German National Library will focus on the following areas in the near future

- | scientific e-journals,
- | other digital journal articles,
- | digital music resources, and
- | the continuation of webpage collections based on the current criteria.

On the other hand, the German National Library intends to reduce the completeness level for digitised works.

Basically, digital copies of publications and other media works from Germany from 1913 are to be collected which, in their analogue form, also fall within the collection mandate of the German National Library. The current collection guidelines stipulate that all digital copies of works originally released in printed form and now published on the Web fall within the collection remit of the German National Library. This section of the collective guidelines needs to be expressed in a more differentiated form. Many digital copies are thus already stored in a stable form and publicly offered for free by the institution which produced the digital secondary form, for example, the one million digital copies of the copyright-free holdings of the Bayerische Staatsbibliothek which were produced in collaboration with Google. The German National Library will not collect and archive these digital copies. It currently plans to proceed as follows:

- | The German National Library will not collect and archive digital copies of genuine archive or museum holdings. It will only do so if the digitised work in its physical form falls within the collection mandate of the German National Library. Responsibility for the digitised material not collected by the German National Library should lie with the institutions under whose collection mandate the original versions fall, or with other trustworthy institutions that fulfil the corresponding tasks for the archive or museum sector. The German National Library fulfils its national bibliographic role here through the Deutsche Digitale Bibliothek (DDB) by contributing its expertise and know-how, thereby ensuring that national bibliographic indexing is conducted. At present, the German National Library assumes the task of coordinating the competence network for the joint DDB project, thus providing the central project management for the further development of the technical infrastructure and the DDB portal.¹⁷ Through these activities, the DDB partners combine virtually the digital

¹⁶Since the amendment of the law, numerous legal reviews have been conducted on individual issues in order to include new aspects and developments in the deliberations of the German National Library. A new data protection and copyright review is to be conducted with regard to the domain crawl.

¹⁷They are responsible for the project management, including coordination of the partners in the DDB competence network, of the technical operator and the implementation partners. The German National Library also operates the DDB's central service point which advises cultural and scientific institutions on organisational and technical questions regarding data delivery and ensures efficient data clearing procedures. The service

collections distributed throughout Germany in archives, libraries, museums and other cultural institutions via a metadata platform. Digitised collections can thus be identified and reused at a single point, even if the individual objects are not archived centrally. Coordination and organisation of the documentation are considered national bibliographic tasks; the archiving tasks are distributed among the cooperation partners.

- | The same applies to digital copies of library holdings published before 1913 which do not fall within the collection mandate of the German National Library and which are indexed by the DDB. The German National Library dispenses with the digitisation of publications published in Germany before 1913 in cases where the digitising bodies are institutions with trusted long-term archives (from experience, these include many state and publicly funded institutions, e.g. the Federal Archives, state libraries or other DINI-certified institutions, i.e. institutions that themselves collect and preserve digital content).
- | Notated music is to be dealt with similarly.
- | For digitised versions of German-language publications that have been published and archived abroad by the respective national library or other trustworthy institution, the German National Library considers that a link is sufficient under the aforementioned conditions (collected by a trustworthy institution with a trusted long-term archive).
- | All digital copies of carrier-bound sound recordings belonging to the collection mandate must be collected independently of the date of origin of the physical work, since there are no comparable institutions with a regional or national collection mandate, and the beginning of the production of sound recording only predated 1913 (when the German National Library began its collection) by a small margin.

Not all institutions are technically and financially able to develop and permanently operate storage, archiving and access modalities for digital publications, or they have other priorities and prefer to use external service providers. Only a few institutions are currently deemed to be suitable for such services in Germany at the federal and the regional (Länder) level. On the other hand, current discussions (e.g. the stabilisation of DDB financing, increase of resources, organisational autonomy) show how difficult it is to assign a cross-border coordinator role. Nonetheless, such constellations must continue to be considered and discussed intensively at the political and institutional level. It must be the goal of the consultations with other legal deposit libraries to arrive at concrete arrangements for cooperative collections of web resources.

For its part, the German National Library (with its archiving service Aredo) offers itself as a trustworthy partner for institutions that cannot themselves build and maintain a trustworthy long-term archive for digital publications. These institutions could provide links to their digital holdings preserved in the German National Library if e.g. there are no other local or regional options or if they should not be used. As part of its strategic considerations, the German National Library will examine the extent to which such a cooperation model can be organised together with similar service providers (such as Aredo) for distributed long-term archiving and availability. The individual regional legal deposit libraries in the German federal states (Länder) represent possible¹⁸ partners in such a model, with the national libraries of other states, especially within the EU, playing an equivalent role for "cross-border publications". The German National Library will help ensure that corresponding bilateral undertakings or arrangements between the Länder or the European states are developed, the distribution of tasks defined and agreement obtained on access models for the

centre is supported by specialised bodies that operate from within the different types of cultural institutions and to which the service centre transfers work packages - in some cases also in the context of formal work contracts or grants.

¹⁸ See appendix: "Stand der Pflichtablieferung von Netzpublikationen in den einzelnen Bundesländern" (=Status of legal deposit of online publications in the individual Länder)

digital media. Partners which are deemed especially suitable here are those with a sound financial, organisational and technical structure.

Examples of "cross-border publications" are the digital music and journal sectors. The German label code has hitherto been a key criterion in structuring the collection of sound recordings. Now that this is no longer used for growing parts of the music market in Germany, it is necessary to look for alternative codes or labels that permit the identification of legal deposit objects across new distribution channels, platforms or streaming services. This question has already been touched on in the "Musik im Netz" workshop. Further discussions are necessary.

It is important to clarify the unresolved issues regarding publication or technical processes and, on this basis, to reach agreement between national libraries on who is responsible for the collection and long-term archiving of certain publications and which other national libraries may access it after the conclusion of relevant contracts.

The same applies to the digital periodicals of international publishers in the fields of science, technology and medicine (STM). Cooperation between national libraries makes obvious sense at the international (i.e. cross-border) level. From a purely economic point of view, central solutions for storage and access are worth considering within Germany. However, there are also good arguments for a cooperative approach based on a number of strong partners (see above).

Possible changes in physical collection

The German National Library has assessed the degree of fulfilment of its collection mandate and questioned the formal collection criteria at regular intervals. The current version of the collection criteria is documented in the collection guidelines.¹⁹

The latest Legal Deposit Regulation dispenses with the deposit and collection of physical editions of daily newspapers, dissertations and publishing-on-demand publications in cases where a digital version is available. This rule should also be adopted for other document groups, where applicable.

In the discussion regarding the collection of digital publications, the fundamental question remains of whether the holdings of the German National Library and the indexing in the Deutsche Nationalbibliografie should be directed at the level of the "work" or at every individual manifestation of the work. At its meeting on 5 December 2016 the Administrative Council of the German National Library rejected the frequently proposed suggestion of reducing the number of physical legal deposit copies to one instead of two physical copies in the case of fully parallel physical and digital issues: *"The Administrative Council [declares itself in favour] of maintaining the collection of two physical copies of a work, even if a digital edition is published in parallel."*

The German National Library currently sees no further possibilities for narrowing its collection mandate. For example, one study concluded that raising the print run figure or the number of pages in the collection guidelines would have little effect: if the minimum number of pages was raised to ten instead of four, then the numbers of printed monographic units collected by the German National Library – 141,841 in 2013 and 137,318 in 2014 – would merely have been reduced by 794 and 640 respectively. An increase in the print-run per edition from 25 e.g. to 50 copies would yield similarly low results; in the age of digital printing, it could even lead to the axeing of scientific and other publications that belong without question to the collection of the German National Library.

¹⁹ Sammelrichtlinien, Stand 1. Mai 2014 (=Collection guidelines, as of 1 May 2014)
<http://d-nb.info/1051940788/34>

A supplement to the collection mandate of the German National Library that includes currently excluded materials, such as analogue and computer games, was discussed within the German National Library and rejected.

Role of the National Library in the digital age

Just as scientists carrying out research into historical phenomena rely today on the physical handing down of scientifically valuable artefacts from earlier times in order to reconstruct a picture of the period in question, it will be important for future generations to have digital evidence in order to obtain a complete picture of our time.

As a national library tasked with preserving the national cultural heritage, the German National Library must collect items authentically without judging the content, document freedom of speech and also defend against external interference. This will enable it to forge a cultural, scientific and sociopolitical image of a nation as a means of creating a bridge between the past, present and future. To what extent, and in which form, it can and must fulfil this role in the digital age is a social and political issue that must repeatedly be discussed and answered.

Anhang

Auszug aus dem Gesetz über die Deutsche Nationalbibliothek (DNBG)

§ 2 Aufgaben, Befugnisse

Die Bibliothek hat die Aufgabe,

1. a) die ab 1913 in Deutschland veröffentlichten Medienwerke und
b) die ab 1913 im Ausland veröffentlichten deutschsprachigen Medienwerke, Übersetzungen deutschsprachiger Medienwerke in andere Sprachen und fremdsprachigen Medienwerke über Deutschland
2. im Original zu sammeln, zu inventarisieren, zu erschließen und bibliografisch zu verzeichnen, auf Dauer zu sichern und für die Allgemeinheit nutzbar zu machen sowie zentrale bibliothekarische und nationalbibliografische Dienste zu leisten,
3. das Deutsche Exilarchiv 1933–1945, die Anne-Frank-Shoah-Bibliothek sowie das Deutsche Buch- und Schriftmuseum zu betreiben,
4. mit den Facheinrichtungen Deutschlands und des Auslands zusammenzuarbeiten sowie in nationalen und internationalen Fachorganisationen mitzuwirken.

§ 3 Medienwerke

- (1) Medienwerke sind alle Darstellungen in Schrift, Bild und Ton, die in körperlicher Form verbreitet oder in unkörperlicher Form der Öffentlichkeit zugänglich gemacht werden.
- (2) Medienwerke in körperlicher Form sind alle Darstellungen auf Papier, elektronischen Datenträgern und anderen Trägern.
- (3) Medienwerke in unkörperlicher Form sind alle Darstellungen in öffentlichen Netzen.
- (4) Filmwerke, bei denen nicht die Musik im Vordergrund steht, sowie ausschließlich im Rundfunk gesendete Werke unterliegen nicht den Bestimmungen dieses Gesetzes.

§ 14 Ablieferungspflicht

- (1) Die Ablieferungspflichtigen haben Medienwerke in körperlicher Form nach § 2 Nr. 1 Buchstabe a in zweifacher Ausfertigung gemäß § 16 Satz 1 abzuliefern. Musiknoten, die lediglich verliehen oder vermietet werden (Miet- oder Leihmateriale), haben die Ablieferungspflichtigen in einfacher Ausfertigung gemäß § 16 Satz 1 abzuliefern.
- (2) Die Ablieferungspflichtigen haben Medienwerke nach § 2 Nr. 1 Buchstabe b in einfacher Ausfertigung gemäß § 16 Satz 1 abzuliefern, wenn eine Inhaberin oder ein Inhaber des ursprünglichen Verbreitungsrechts den Sitz, eine Betriebsstätte oder den Hauptwohnsitz in Deutschland hat.
- (3) Die Ablieferungspflichtigen haben Medienwerke in unkörperlicher Form nach § 2 Nr. 1 Buchstabe a in einfacher Ausfertigung gemäß § 16 Satz 1 abzuliefern.
- (4) Wird die Ablieferungspflicht nicht binnen einer Woche seit Beginn der Verbreitung oder der öffentlichen Zugänglichmachung des Medienwerkes erfüllt, ist die Bibliothek nach Mahnung und fruchtlosem Ablauf von weiteren drei Wochen berechtigt, die Medienwerke auf Kosten der Ablieferungspflichtigen anderweitig zu beschaffen.

§ 15 Ablieferungspflichtige

Ablieferungspflichtig ist, wer berechtigt ist, das Medienwerk zu verbreiten oder öffentlich zugänglich zu machen und den Sitz, eine Betriebsstätte oder den Hauptwohnsitz in Deutschland hat.

§ 16 Ablieferungsverfahren

Die Ablieferungspflichtigen haben die Medienwerke vollständig, in einwandfreiem, nicht befristet benutzbarem Zustand und zur dauerhaften Archivierung durch die Bibliothek geeignet unentgeltlich und auf eigene Kosten binnen einer Woche seit Beginn der Verbreitung oder der öffentlichen Zugänglichmachung an die Bibliothek oder der von dieser benannten Stelle abzuliefern.

Medienwerke in unkörperlicher Form können nach den Maßgaben der Bibliothek auch zur Abholung bereitgestellt werden.

§ 20 Verordnungsermächtigung

Zur geordneten Durchführung der Pflichtablieferung und um einen nicht vertretbaren Aufwand der Bibliothek sowie um Unbilligkeiten zu vermeiden, wird das für Kultur und Medien zuständige Mitglied der Bundesregierung ermächtigt, durch Rechtsverordnung zu regeln:

1. die Einschränkung der Ablieferungs- oder der Sammelpflicht für bestimmte Gattungen von Medienwerken, wenn für deren Sammlung, Inventarisierung, Erschließung, Sicherung und Nutzbarmachung kein öffentliches Interesse besteht,
2. die Beschaffenheit der ablieferungspflichtigen Medienwerke und die Ablieferung in Fällen, in denen ein Medienwerk in verschiedenen Ausgaben oder Fassungen verbreitet oder öffentlich zugänglich gemacht wird,
3. das Verfahren der Ablieferung der Medienwerke sowie
4. die Voraussetzungen und das Verfahren bei der Gewährung von Zuschüssen.

Auszug aus der Verordnung über die Pflichtablieferung von Medienwerken an die Deutsche Nationalbibliothek (Pflichtablieferungsverordnung – PflAV)

§ 1 Einschränkung der Ablieferungspflicht

- (1) Zur Erfüllung der Aufgaben der Deutschen Nationalbibliothek (Bibliothek) sind Medienwerke von den Ablieferungspflichtigen nach den Maßgaben der §§ 14 bis 16 des Gesetzes über die Deutsche Nationalbibliothek an die Bibliothek abzuliefern, soweit sich aus dieser Verordnung nichts anderes ergibt. Unbeschadet der §§ 3, 4, 8 und 9 kann die Bibliothek auf die Ablieferung verzichten, wenn an der Sammlung kein öffentliches Interesse besteht.
- (2) Ein Anspruch auf Aufnahme eines Medienwerkes in die Sammlung der Bibliothek besteht nicht.

§ 7 Beschaffenheit von Netzpublikationen und Umfang der Ablieferungspflicht

- (1) Unkörperliche Medienwerke (Netzpublikationen) sind in marktüblicher Ausführung und in mit marktüblichen Hilfsmitteln benutzbarem Zustand abzuliefern. Eine Pflicht zur Ablieferung besteht nicht, wenn die Ablieferungspflichtigen im Rahmen des § 16 Satz 2 des Gesetzes über die Deutsche Nationalbibliothek mit der Bibliothek vereinbaren, die Netzpublikationen zur elektronischen Abholung bereitzustellen. Für die Ablieferung von Netzpublikationen gilt § 2 Abs. 3 entsprechend; für die Bereitstellung zur elektronischen Abholung gilt § 2 Abs. 3 Satz 1 entsprechend.
- (2) Die Ablieferungspflicht umfasst auch alle Elemente, Software und Werkzeuge, die in physischer oder in elektronischer Form erkennbar zu den ablieferungspflichtigen Netzpublikationen gehören, auch wenn sie für sich allein nicht der Ablieferungspflicht unterliegen. Dies gilt insbesondere für nicht marktübliche Hilfsmittel, die eine Bereitstellung und Benutzung der Netzpublikationen erst ermöglichen und bei den Ablieferungspflichtigen erschienen sind. Sie sind zusammen mit den Netzpublikationen abzuliefern oder zur elektronischen Abholung bereitzustellen.

§ 8 Einschränkung der Ablieferungspflicht für Netzpublikationen in verschiedenen Ausgaben und aufgrund technischer Verfahren

- (1) Die Bibliothek kann auf die Ablieferung oder elektronische Abholung einzelner Ausgaben von Netzpublikationen verzichten, wenn diese gleichzeitig oder nacheinander in unterschiedlichen technischen Ausführungen erscheinen.
- (2) Die Bibliothek kann auf die Ablieferung verzichten, wenn technische Verfahren die Sammlung und Archivierung nicht oder nur mit beträchtlichem Aufwand erlauben. Sie kann nicht sammelpflichtige Netzpublikationen archivieren, wenn zur Sammlung eingesetzte automatisierte Verfahren eine Aussonderung solcher Netzpublikationen nicht oder nur mit beträchtlichem Aufwand erlauben.
- (3) Umfang und Häufigkeit der Ablieferung von regelmäßig aktualisierten Netzpublikationen können durch die Bibliothek eingeschränkt werden.

§ 9 Weitere Einschränkungen der Ablieferungspflicht für Netzpublikationen

Nicht abzuliefern sind

1. Netzpublikationen, die den in § 4 Nr. 8, 10, 13 und 14 bezeichneten Medienwerken entsprechen, sowie lediglich privaten Zwecken dienende Websites,
2. zeitlich befristete unkörperliche Vorab- und Demonstrationsversionen zu körperlichen oder unkörperlichen Medienwerken, sofern sie nach Erscheinen der endgültigen Publikation wieder aus dem Netz genommen werden,
3. selbstständig veröffentlichte Betriebssysteme, nicht sachbezogene Anwenderprogramme, die nicht unter § 7 Abs. 2 fallen, sachbezogene Anwendungswerkzeuge zur Nutzung bestimmter Internetdienste, Arbeits- und Verfahrensbeschreibungen,
4. Bestandsverzeichnisse, soweit sie nicht von einem Dritten veröffentlicht werden,
5. Netzpublikationen, die aus Fernseh- und Hörfunkproduktionen abgeleitet werden, soweit sie nicht von einem Dritten veröffentlicht werden,
6. inhaltlich unveränderte Spiegelungen von Netzpublikationen, soweit die ursprüngliche Veröffentlichung abgeliefert wurde,
7. netzbasierte Kommunikations-, Diskussions- oder Informationsinstrumente ohne sachliche oder personenbezogene Zusammenhänge,
8. E-Mail-Newsletter ohne Webarchiv,
9. Netzpublikationen, die nur einer privaten Nutzergruppe zugänglich gemacht sind,
10. selbstständig veröffentlichte Primär-, Forschungs- und Rohdaten.

Auszug aus der Begründung zur Gesetzesnovellierung

Zu § 2 (Aufgaben, Befugnisse)

...

Im Bereich der Printmedien gibt es überwiegend Publikationen von Verlagen sowie von wissenschaftlichen und kulturellen Einrichtungen. Damit besteht ein Großteil der derzeitigen Sammlung der Bibliothek aus solchen Veröffentlichungen. Darüber hinaus werden bereits heute Netzpublikationen der genannten Stellen auf freiwilliger Basis von der Bibliothek gesammelt. Mit der Einbeziehung dieser Publikationsformen in den Sammelauftrag der Bibliothek wird dies auf eine gesetzliche Grundlage gestellt. Damit wird der technologischen Entwicklung und der sich wandelnden und an Bedeutung zunehmenden Form der Veröffentlichung gefolgt.

In der modernen Entwicklung der Netzpublikationen wächst demgegenüber die Zahl der Veröffentlichungen aus dem nicht gewerblichen und nicht institutionellen Bereich sehr stark. Die daraus entstehende Fülle von einzelnen digitalen Objekten kann aus wirtschaftlichen Gründen nicht in einer gezielten Auswahl gesammelt werden. Vielmehr sollen für diesen Sammlungszweig automatische Such- und Speichermethoden für das umfassende Sammeln von Internetseiten zum Einsatz kommen (Web-Harvesting). Hierbei werden die Inhalte von bestimmten Bereichen (zum Beispiel alle Websites mit der Domain „.de“) automatisiert eingesammelt. Dabei kann nicht jede Information zu jedem Zeitpunkt gesichert und erschlossen werden. Vielmehr sollen beispielsweise zwei Mal jährlich zeitliche Schnitte durch diesen bestimmten Bereich des Internets gelegt werden. Auf diese Art kann dokumentiert und archiviert werden, welche Information zu einem bestimmten Zeitpunkt im Internet vorhanden war. Eine intellektuelle Erschließung dieser Dokumente soll nicht stattfinden. Einzelheiten werden in der Rechtsverordnung nach § 20 dieses Gesetzes geregelt.

Neben dem Anstieg der unkörperlichen Publikationen in den genannten traditionellen Sammelbereichen der Bibliothek ist von der sprunghaft steigenden Zahl unkörperlicher Veröffentlichungen insbesondere der nicht kommerziell wissenschaftliche wie populäre Bereich betroffen. Diese Publikationen prägen ein größer werdendes Feld der Kultur Deutschlands. Diese in das Internet eingestellten Veröffentlichungen betreffen auf unterschiedlichen Qualitätsniveaus alle Wissensbereiche. Deren dauerhafte Verfügbarkeit zu sichern – wie dies bereits seit Inkrafttreten des noch geltenden Gesetzes für die körperliche Form erfolgt – ist für eine Kulturnation unverzichtbar. Andernfalls würde es künftig ausgeschlossen sein, Entwicklungen zu beschreiben, zu nutzen und in Ursache und Wirkung zu analysieren. Die Bereitstellung des Materials dafür ist international anerkannte Aufgabe einer Nationalbibliothek.

Die Menge ständig neu erscheinender digitaler Publikationen im Internet (geschätzt werden 2004 über sieben Millionen Websites allein mit der Top-Level-Domain „.de“ mit einer Vielzahl dort auftauchender Einzelveröffentlichungen) schließt konventionelle Methoden und Mittel für die Sammlung, Archivierung und Bereitstellung aus. Eine regelmäßige, stetige Suche, Identifizierung, intellektuelle Erschließung oder eine Sichtung nach Bedeutung ist insbesondere aus finanziellen Gründen ausgeschlossen. Die Identifikation und Selektion sammelpflichtiger Publikationen muss nach „maschinenlesbaren“ Kriterien erfolgen. Nur so kann die Wirtschaftlichkeit der Aufgabenerfüllung im Rahmen der verfügbaren Haushaltsmittel gewährleistet werden.

Damit lassen sich zwei „Säulen“ der Sammlung elektronischer Publikationen charakterisieren: zum einen die gezielte Sammlung von Verlagsveröffentlichungen, wissenschaftlichen, institutionellen und kulturellen Publikationen mit den ihnen eigenen Qualitätsreferenzen aufgrund eigens gewährter Zugänge zu Websites oder in besonderen Ablieferungsverfahren; zum anderen die stichprobenweise Sammlung von Internetquellen über Harvesting-Verfahren.

Um nicht kommerzielle Veröffentlichungen, die so genannte graue Literatur, deutlicher als in der bisherigen Gesetzesfassung in den Sammelauftrag einzubeziehen, wird (ähnlich den Schweizer Regelungen vom 18. Dezember 1992) umfassend auf „veröffentlichte“ Medienwerke abgestellt. Die in § 2 Abs. 1 des noch geltenden Gesetzes verwendeten Begriffe „verlegt“ oder „hergestellt“ bedurften einer eigenen Definition in § 19 Abs. 2 des noch geltenden Gesetzes, die beispielsweise auch nicht kommerzielle Selbstverleger gesondert erwähnte, um nicht als allein auf Gewerbebetriebe zielende Kriterien missverstanden zu werden.

Nummer 1a definiert den Sammelauftrag für die im Geltungsbereich des Gesetzes veröffentlichten Medienwerke, Nummer 1b für die außerhalb des Geltungsbereiches veröffentlichten Medienwerke. Der Sammelauftrag leitet sich vom Grundanliegen der Bibliothek ab und ist Voraussetzung für die Umsetzung aller weiteren Aufgaben. Der Sammlungsbeginn ab 1913 ist darauf zurückzuführen, dass 1912 am Standort Leipzig die Deutsche Bücherei als Archiv des deutschen Schrifttums gegründet wurde. Aufgrund der Zusammenführung der Häuser in Leipzig und Frankfurt am Main infolge der Wiedervereinigung Deutschlands wurden die im Wesentlichen übereinstimmenden Sammelaufgaben abgestimmt und vereinheitlicht, so dass heute der Bestandsaufbau und die Erschließung arbeitsteilig nach gemeinsam erarbeiteten Richtlinien erfolgen. Sie unterliegen ständigen Präzisierungen und Konkretisierungen, um den Veränderungen im Publikationsbereich und den Bedürfnissen der Nutzer Rechnung zu tragen.

Die Einbeziehung der Übersetzungen deutschsprachiger Autoren in fremde Sprachen und fremdsprachiger Veröffentlichungen über Deutschland und Personen des deutschen Sprachgebietes entspricht § 2 Abs. 1 Nr. 2 des noch geltenden Gesetzes. Die Bestimmung trägt dem Aspekt Rechnung, die Rezeption der deutschen Sprache und Kultur im Ausland zu belegen. Sie dient dem geistigen Austausch zwischen den Völkern in einer immer enger werdenden kulturellen Verflechtung.

Stand der Pflichtablieferung von Netzpublikationen in den einzelnen Bundesländern

(Stand: 2. November 2016)

Baden-Württemberg	gesetzliche Regelung seit 01.01.2007
Bayern	Ablieferung nur auf freiwilliger Basis
Berlin	keine Ablieferung
Brandenburg	gesetzliche Regelung seit 21.06.2012
Bremen	keine Ablieferung
Hamburg	gesetzliche Regelung seit 08.09.2009
Hessen	gesetzliche Regelung seit 27.11.2012
Mecklenburg-Vorpommern	keine Ablieferung
Niedersachsen	gesetzliche Regelung nur für digitale Amtsdruckschriften
Nordrhein-Westfalen	gesetzliche Regelung seit 01.01.2013
Rheinland-Pfalz	gesetzliche Regelung seit 03.12.2014
Saarland	gesetzliche Regelung seit 01.12.2015
Sachsen	gesetzliche Regelung seit 01.01.2014
Sachsen-Anhalt	gesetzliche Regelung seit 02.05.2013
Schleswig-Holstein	gesetzliche Regelung seit 30.08.2016
Thüringen	gesetzliche Regelung seit 08.02.2011