

2016 Satellite meeting - *News, new roles & preservation advocacy: moving libraries into action*
10 – 11 August 2016
Lexington, Kentucky USA, USA

“At Risk”: An Analysis of newspaper source materials digitized by U.S. and European repositories

Maria Smith

Digital Repository Analyst, International Resources, Center for Research Libraries, Chicago, United States.

E-mail address: msmith@crl.edu



Copyright © 2016 by **Maria Smith**. This work is made available under the terms of the Creative Commons Attribution 4.0 Unported License: <https://creativecommons.org/licenses/by/4.0/>

Abstract:

In 2015, in cooperation with IFLA News Media Section, The Center for Research Libraries conducted an in-depth assessment of the outputs of the major newspaper digitization efforts in the U.S. and Europe. The purpose of the assessment was to identify the comparative strengths and gaps in the corpus of newspapers digitized to date, relative to the total body of newspapers held by libraries. The findings of that assessment revealed a number of areas where cooperative action was required by libraries to address the persistent limitations in our shared effort.

A key observation of that assessment was that the bulk of newspaper digitization has been conducted from microfilm surrogates, meaning that an undetermined—but likely significant—quantity of titles and issues remain at risk due to fragility, lack of preservation, damage from use, displacement, or sheer neglect. Through a follow up assessment, CRL has attempted to better determine the scale of the “at risk” corpus by examining the format of source materials used in digitization of European and American newspapers by national libraries, regional collecting institutions, and commercial providers.

This paper presents results of CRL’s assessment. Based on quantitative and qualitative information, it attempts to quantify the number of newspaper titles digitally reproduced by the major repositories using original print copies compared to the amount of content scanned from microfilm. The author characterizes libraries’ varied approaches to newspaper digitization, particularly from the perspective of content sourcing, and the implications of these choices on the future accessibility of content and the survival of the historical print record.

Strategies for newspaper digitization are often motivated by efficient access to as much content as possible. However, this paper argues that preservation decisions must also play a role in selection, as

original print copies never selected to be microfilmed are continuing to be neglected for preservation measures.

Keywords: Newspapers, preservation, digitization, at-risk content

Introduction

From 2014 to 2106, the Center for Research Libraries (CRL) through the International Coalition on Newspapers (ICON) conducted a study to evaluate the effectiveness of major efforts undertaken by libraries and electronic publishers to provide sustainable electronic access to newspaper back files. Sponsored in part by a grant from the Andrew W. Mellon Foundation, CRL's assessment sought to measure the current scope and coverage of historical world newspaper digitization underway, and to evaluate library and publisher operations, systems and practices to gauge the reliability of those practices in ensuring long-term accessibility and sustainability of news resources.

The impetus for the assessment grew out of libraries' increasing reliance on digital surrogates of historical newspaper content for access and development of research collections. Academic libraries and many national libraries invest considerable sums to digitize newspapers to make them more accessible to historians and other researchers. And each year, research libraries in the aggregate spend millions of dollars to purchase from commercial publishers databases of digitized historical newspapers. At the same time, many libraries are under growing pressure to trim collections and reduce storage of bulky and increasingly embrittled collections of newsprint. As a result, libraries must make consequential decisions daily about whether to preserve, conserve, reformat, and even retain these types of important materials. The two primary purposes of the assessment were: 1) to inform library decisions about the purchase and licensing of newspaper databases, and 2) to support strategic library and publisher investment in development, implementation and management of digital newspapers programs.

This paper summarizes three major outputs of the CRL/ICON assessment that may inform future, strategic newspaper digitization initiatives. Those outputs are:

1. A comparative analysis of European & American newspaper digitization to date, relative to the total body of newspapers held by libraries, presented at the ICON Summit in April, 2015.¹
2. The percentage breakdown of source material format types used in the digitization of newspapers by European, British and American libraries.
3. The approximate number of titles from several North American and European countries and repositories that remain available only in print format, informing a projected estimate on the total body of newspaper materials that remain at risk worldwide.

The data presented herein are based on information gathered by CRL on major newspaper digitization efforts by national and academic libraries, major producers of commercial academic producers, and other significant stakeholders in the newspaper digitization arena.

¹ "Framing a Common Agenda for Newspaper Digitization and Preservation: An ICON Summit," <http://www.crl.edu/events/framing-common-agenda-newspaper-digitization-and-preservation-icon-summit> (accessed June 27, 2016)

Part 1: Comparative analysis of European & American newspaper digitization to date

In 2015, in order to identify the comparative strengths and gaps in the corpus of newspapers digitized to date, CRL conducted an assessment of the total output of newspaper digitization. The assessment used title-level information available from major newspaper digitization projects, and issue-level information aggregated by CRL in the ICON Database of International Newspapers (<http://icon.crl.edu>). The ICON Database is a registry of holdings information on the hard copy, microform and digitized holdings of U.S. and foreign newspapers. CRL's assessment of global newspaper digitization examined nearly 12,000 titles from approximately 90 different digital collections

Based on information gathered by CRL, supplemented by reports from major initiatives such as Europeana, a conservative estimate of the total number of newspapers from all world regions digitized to date (any years) exceeds 30,000 titles. Given that the research was narrowly confined to digitized newspapers held in major repositories in the U.S., U.K. and Europe, the number of digitized titles available worldwide could easily exceed 45,000 when all countries are taken into account.

A survey conducted by Europeana Newspapers in 2012 found that the number of digitized newspapers in Europe was estimated at 24,000 titles, with over 129 million pages scanned as of 2012. CRL's examination of newspaper coverage from Western Europe through the major national library and/or regional academic efforts revealed that peak coverage begins in the 1880's and runs through 1918 (likely a product of the European-wide emphasis on digitizing newspapers from the World War I era). Recent projects initiated to scan newspapers from the World War II period produce a spike in coverage from 1944-1945.

The 2012 Europeana Newspapers survey suggested that few libraries had digitized more than 10% of their collection (either in terms of titles or page numbers). Libraries in Europe have taken varying approaches to digitization, ranging from thematic selection, chronological coverage, and historical significance to comprehensive coverage of cultural heritage from the earliest available material forward. Though exact numbers of newspapers published in Europe are difficult to ascertain, it may be assumed that there is significant room for growth in the overall coverage of newspaper digitization in Europe.

Compared to European efforts, newspaper digitization in the United States is less deep in terms of chronological coverage and more diffuse (in terms of stakeholders and access models). Library participation in digitization began in earnest with the federally-funded National Digital Newspapers Program (NDNP) initiated in 2004. Coordinated by the Library of Congress, and funded by the National Endowment for the Humanities, the NDNP has awarded grants to institutions in 39 states and territories to date, for digitization of newspapers published during the period 1836–1922.

As of April 6, 2014, 1,728 digitized newspaper titles were represented in *Chronicling America*, the Library of Congress' platform for access to the NDNP newspapers.² As funding and scope for the NDNP is limited, many of the current and former participants have continued digitizing content beyond the scope of the NDNP. A 2014 survey of 35 current and former participants in the program—academic libraries, state historical societies, and state libraries—revealed that these institutions had produced to date more than 15 million pages of content beyond the scope of NDNP. CRL's coverage assessment included 10

² As of June 27, 2016, the number of titles represented in *Chronicling America* had increased to 2,088.

current or former participants of the NDNP that digitize newspapers beyond the scope of NDNP. From this sampling, 1,250 titles “beyond NDNP” were identified, representing publications dating from as early as 1787 and as recent as the present day.

In the United States, numerous commercial providers have also been actively digitizing and hosting newspaper content, in databases made available mostly to academic and public libraries through one-time purchase or ongoing annual subscriptions. Major providers include East View Information Services, Gale Cengage, ProQuest, and Readex (a division of NewsBank). CRL’s assessment included multiple collections from these providers (whether single-title products or historical news aggregations), identifying 2,212 additional titles.

While data gathered by CRL to date covers only a representative sampling of U.S. titles digitized, even a generous tripling of the estimated number of titles digitized would suggest that the U.S. has not yet reached 10% of total coverage of pre-1923 newspapers held by libraries. For instance, based on the number of titles cataloged through the U.S. Newspapers program, it is estimated that in 1922 there were more than 15,000 titles currently being published in the United States. CRL’s sampling for 1922 indicated only 360 titles available in digitized format. Even at the peak of digitized news coverage (1908) in the sample, the number of titles represented was less than 4% of the total press output in the U.S.

With the data CRL has gathered, it is difficult to estimate with any precision the amount of newspaper content digitized relative to the total number of newspapers held by libraries worldwide. However, CRL’s sampling indicates that the percentage of digitized content compared to the total number of newspapers is quite small. Many digitization projects are working from microfilm “vaults,” which leaves the most at-risk material unavailable. Titles that have not been scanned or otherwise preserved remain “at risk” due to the highly acidic nature of much newsprint, potential damage through circulation and handling, and sheer neglect.

Part 2: European and American newspaper digitization source materials

To further characterize library approaches to newspaper digitization in terms of preservation risks, CRL has identified the approximate number of newspaper titles that were digitally reproduced using original print copies as compared to the amount of content that has been scanned from microfilm by American, British and European repositories. These figures are based on data gathered from electronic newspaper repositories, library catalogs, presentations, articles and reports, and consist of qualitative and quantitative information on source materials utilized by European and American repositories for digitized content.

Data aggregation

CRL’s process of identifying the source materials of digitized newspapers held by European and American repositories consisted of aggregating descriptive-qualitative and quantitative data. CRL examined digitized content from the major European and American newspaper databases included in the original coverage assessment (a complete list of those collecting institutions is appended to this document). The first step in the evaluation was an investigation of publicly available information and anecdotal evidence on digitization processes from library or product websites and from articles or presentations published by or about the various repositories. This information was used to identify or

confirm the use of print or microform materials and further characterize digitization processes and outcomes through evidence found in collection and product descriptions, digitization policies, workflow documentation, and descriptive or administrative metadata.

CRL then conducted a sampling of digitized page images from the online newspaper repositories to verify whether pages were scanned from original print sources or through microform conversion.³ Visual cues were used to identify the object source: full-color images suggested that the digital images had been produced using original paper materials, while grayscale or bi-tonal images suggest that the source materials may be derived microfilm copies. To account for the fact that digital versions may have been produced using paper originals while still producing bi-tonal or grayscale images, CRL used other visual cues to indicate use of microfilm copies in digitization such as: visible scratches, graininess, and stark black borders around the page image. Where available, metadata was reviewed to confirm the source material of the digital copy: however, of the 57 newspaper databases surveyed, only 4 institutions have exposed metadata in their access platform that captures provenance information on the source materials used for digitization. Given these limitations, the title counts found below should be considered an informed estimate of the amount of digitized content created using print copies and not a confirmed total.

Europe

Of the 46 digital repositories surveyed from European countries, CRL found that 20 repositories have sourced at least a portion of their digitized newspaper collections from print holdings. European repositories that have been confirmed through quantitative and (in most cases) qualitative evidence to have sourced all or some of their digitized newspapers from print holdings include:

1. Austrian National Library
2. Biblioteca comunale dell'Archiginnasio
3. Biblioteca Panizzi Reggio Emilia
4. Biblioteca Universitaria di Pisa
5. Bibliothèque cantonale et universitaire – Lausanne
6. Bibliothèque nationale de France
7. Bibliothèque royale de Belgique
8. British Library / British Newspaper Archive, Ltd.
9. Dr. Friedrich Teßmann Library
10. Hamburger Abenblatt
11. Koninklijke Bibliotheek
12. National Library of Norway
13. National Library of Sweden
14. The National Library of Wales
15. Staatsbibliothek zu Berlin
16. Swiss National Library
17. Trinity News Archive
18. Universitäts- und Landesbibliothek Bonn
19. Universitäts- und Landesbibliothek Darmstadt
20. Universitätsbibliothek Heidelberg

³ The sampling did not review every issue of the newspaper titles surveyed; therefore a given title may have been digitized from both microfilm and paper, but only one or the other is reflected in the title counts found on page 6.

The lack of granular issue-level data from European institutions on the source or provenance of digital imaging restricts CRL’s ability to estimate with any precision the amount of newspaper content digitized from print versus microfilm. Thus, CRL’s analysis is limited to title-level information.

Assessing the European “corpus” of available titles, approximately 40%—2,377 titles—have been digitized, at least in part, from print holdings. However, this overall statistic may be misleading. Of the 20 organizations, only three (The Austrian National Library, The National Library of Wales, and Koninklijke Bibliotheek) reported sourcing the majority or all of their digitized content from print holdings. Breaking the estimated percentage of print titles down by country shows that most other programs rely heavily on microfilm. Factoring ANNO, Wales and the KB out of the total percentage, the remaining institutions have collectively sourced approximately 6% of all titles from print originals.

Location/Country	Total #of titles	# of titles digitized from print	Percentage of print sourced titles out of total
Europe (total)	5786	2377	41%
Austria*	799	799	100%
Belgium	80	5	6.3%
Denmark*	15	0	0%
England	976	57	5.8%
Finland*	342	0	0%
France	47	4	8.5%
Germany	375	17	4.5%
Ireland	137	3	2.2%
Isle of Man	31	0	0%
Italy	882	13	1.5%
Netherlands*	1398	1275	91.2%
Norway*	391	66	16.9%
Scotland	80	19	23.8%
Sweden	24	6	25%
Switzerland*	77	7	9.1%
Wales*	135	107	79.3%

*Indicates that title count has been confirmed by holding repository through email, or Library publication/website

Weighing the percentage of print sources compared to microfilm over time, the aggregated data shows that coverage for early years (1618-1719) was largely digitized from print holdings. However digitized content from succeeding publication years (1720-2015) was sourced overwhelmingly from microfilm copies (75% from microfilm, on average). The following chart demonstrates the percentages of European content digitized from original print copies, broken down by 50-year time spans:

Publication Year Span	1618-1668	1669-1719	1720-1770	1771-1821	1822-1872	1873-1923	1924-1974	1975-2015
Avg. %	98%	51%	28%	31%	24%	26%	26%	11%

While titles digitized from print make up approximately 40% of the total count of European titles digitized, the **volume** of content digitized from print copies is actually significantly less than 40% of all content that is digitally available. Many titles digitized from print are often shorter in chronological span, ranging from a few issues to scattered years.

United States

To an even greater extent than European projects, newspaper digitization in the United States has relied heavily on microfilm holdings to produce digital collections of historical newspaper content. According to CRL's findings, of 16 libraries assessed, 10 U.S. libraries have digitized a portion of their online collection from print copies .U.S. repositories that have been confirmed to have sourced some of their digitized newspaper collections from print holdings include:

1. California Digital Newspaper Program
2. Connecticut Digital Newspaper Project
3. Hoosier State Chronicles
4. Illinois Digital Newspaper Collection
5. Indiana State University
6. Kentucky Digital Library
7. Louisiana Newspaper Access Program
8. Montana Memory
9. Texas Digital Newspaper Program
10. Utah Digital Newspapers

These institutions collectively report 143 titles partially or fully scanned from print out of a total 4,734 titles.

The most significant library newspaper digitization accomplishment to date in the United States has been managed by the Library of Congress and the National Endowment for the Humanities (NEH) through the NDNP. The NDNP has facilitated the digitization of more than 2,000 titles from 39 states and territories. Under NDNP, microfilm copies are used for all digital conversion.

Several state and university libraries, currently or previously working under an NDNP grant, have gone on to expand their digitized collections to include additional newspaper content that is outside of the NDNP scope. Under their extended projects, libraries are not limited to using microfilm copies for digitization. However, the majority of U.S. libraries that have extended their digitization beyond the NDNP project have continued to rely on microfilm copies for digital reproduction.

The most significant undertaking of print material digitization in the United States was conducted by the University of Kentucky Libraries. The Kentucky Digital Library expanded the coverage of publications to include 126 titles not represented in the NDNP collection. Of the content the University of Kentucky has independently digitized, 67 titles were sourced from print holdings, which is approximately 53% of their non-NDNP collection. However, most other institutions that have digitized from print have done so only to fill in small gaps in the microfilm coverage or to replace low quality microfilm. Overwhelmingly, digitization in the United States continues to rely on microfilm holdings, with 96% of all digital newspaper collections at libraries sourced from micro-reformatted content.

Part 3: Newspaper materials held exclusively in print

CRL's assessment established that the amount of digitization from print thus far has been small—even insignificant—relative to the amount of scanning from microfilm. Microfilm digitization efforts to date appear to be driven more by economics and ease of access (*i.e.*, from microform) rather than by collection preservation priorities (*i.e.*, from print).

Thus, titles that exist only in print—not microfilmed or otherwise reformatted—remain at considerable risk due to ongoing neglect of these materials in preservation/digitization priorities. CRL sought to provide a measurement of the “at-risk” corpus of print-only material to better understand the scale of the potential threat to this voluminous, acidic and highly fragile material.

The varying consistency and granularity of newspaper records from state and national libraries hindered CRL's attempts to produce a comprehensive assessment of news sources collected and reformatted by libraries and archives. The figures presented below are therefore the result of various methods of information gathering by CRL. CRL consulted national library websites, catalogs, and databases as well as information maintained by collaborative library projects from a single country to ascertain efforts to collect or identify comprehensive newspaper materials that have been published in a specific country or region. Where possible, CRL used catalog or database records to generate a count of title level holdings from surveyed institutions to identify a figure that represents the total number of newspapers that have been published in a single country. CRL then attempted to identify how many of those titles have been microfilmed by holding institutions in order to indicate the amount of print content that remains unpreserved. When the library catalog or database records could not be used to generate total newspaper counts by format, CRL looked to alternative sources to identify an estimated total. For example: professional publications, contacting and gathering information from repository representatives and collection descriptions from library websites.

Aggregating accurate and comprehensive data on the overall body of newspaper materials available worldwide is infeasible, given the inconsistency of library records and lack of authoritative holdings information from most libraries, and the impossibility of incorporating every potential stakeholder. Thus, CRL assessed data where available, solicited input from selected libraries, and conducted case studies to suggest implications for the broader potential corpus of at-risk materials that exists worldwide.

United States

Through the United States Newspaper Program (USNP), the Library of Congress facilitated the creation of the “U.S. Newspaper Directory,” a searchable database that represents newspapers published in the United States since 1690 as well as library holdings.⁴ The directory attempts to represent only distinct title records in the database, with all various formats associated to a single publication record. As of May 2016, the directory lists 153,905 newspaper titles that have been published and cataloged in the U.S.

Of that total, approximately 98,087 titles have associated microform records, of any generation, reported as being held by at least one U.S. library. Acknowledging the limitations of the data (including accuracy, timeliness, and potential omissions), by this measure there are as many as 50,000 titles in the U.S. that are currently held only in print by any library or collecting institution. This implies that approximately

⁴ “U.S. Newspaper Directory, 1690-Present” <http://chroniclingamerica.loc.gov/search/titles/>

32% of all newspapers published in the United States are currently held in print and have not been converted into micro or digital format for preservation or access purposes.

Canada

The Canadian memory and heritage coalition Canadiana is currently carrying out bibliographic research on newspapers that were published in the country up through the year 1925. Canadiana's research is approximately 70% complete. Canadiana has thus far identified 6,807 newspaper titles, and estimates that the total number of titles will reach 7500 or 8000 titles.

Canadiana reports that 75% of these titles are available on microform, and additionally estimates that about 10 to 15% of the microfilmed titles are also available in digital format. Canadiana's findings imply that 25% (approximately 1,700 titles) of the cataloged Canadian newspapers have not been converted into micro- or digital format. Considering the project's limited scope, there likely is a great deal more Canadian newspaper content that remains unidentified and at-risk.⁵

Europe

Through legal deposit, the Bibliothèque nationale de France (BnF) has built a collection of approximately 60,000 historic and current newspaper titles. The library began a program of systematic microfilming in the 1960's, preserving the majority of France's "national-level newspapers as well as numerous regional dailies and local editions. In 2006, the BnF's microfilm collection numbered approximately 2,300 titles of newspaper materials.⁶ BnF itself confirmed that the amount of content preserved on film was insufficient: since its microfilm program began in 1958, only 2.5 linear kilometers of newspaper volumes had been preserved, out of more than 10km of print.⁷ Though the BnF also procured numerous titles on film from outside parties (including several thousand produced by A.C.R.P.P.), the vast majority of French newspaper materials have not been reformatted into microfilm or digital versions (by the figures above, as much as 75% of printed content (volumes) or 96% of titles published).

The National Library of the Netherlands, Koninklijke Bibliotheek, has estimated that there have been more than 7,000 newspaper titles published in the Netherlands since 1618. Only a small selection of Dutch news publications is available in digital or micro formats. According to research conducted by Metamorfoze, a national program in the Netherlands for the preservation of paper heritage, it is estimated that 20% of all Dutch newspapers are available in microfilm.⁸ Additionally, according to an inventory constructed by Metamorfoze, last updated in 2016, there are a total of 1,966 Dutch titles held in either microfilm, digital format or both, implying that there could be as many as 5,034 Dutch news titles (approximately 72%) held only in print format.⁹

In Sweden, the National Library has committed to microfilming all incoming newspaper materials since 1979. The Library has been taking in newspaper publications since 1661, and a portion of pre-1979

⁵ Correspondence with Beth Stover, Collections Manager, Canadiana, April 18 2016

⁶ Delaunay, Else. "Newspaper Digitisation in France: Present Developments." (IFLA publication #122, 2006). [Google Books](#)

⁷ Delaunay/Rollet. "Newspaper Access and Preservation in France." (IFLA publication #118, 2006). [Google Books](#).

⁸ Koninklijke Bibliotheek website "The DDD Project" <https://www.kb.nl/en/organisation/research-expertise/digitization-projects-in-the-kb/databank-of-digital-daily-newspapers/the-ddd-project>

⁹ Metamorfoze "Krantentitels"

https://www.metamorfoze.nl/sites/metamorfoze.nl/files/publicatie_documenten/Kranten%20geconserveerd%20MM%20jan-2016.pdf

content has been selected for microfilm reformatting by the Library. The majority of the newspapers published from 1645 to 1978 are still held exclusively on print.¹⁰ Of the total 5,790 Swedish titles held by the national repository, the Library's catalog records indicate that 1,200 titles are available, at least in part, in microform.¹¹ Therefore, approximately 4,384 titles published between 1661 and 1978 are held exclusively in print by the Library, making up 75% of their total newspaper collection.

In the United Kingdom, the British Library and the network comprising NEWSPLAN (a co-operative program covering all the regions of the United Kingdom and Ireland) worked through the 1980's and 1990's to systematically identify all available newspapers held by regional libraries. Definitive numbers of titles identified appear not to have been aggregated, but an assessment of regional NEWSPLAN reports and available databases suggest that the number of titles identified in England, Ireland, Scotland and Wales exceeded 21,800 titles.¹²

It cannot be determined at this time how many of the newspapers identified under NEWSPLAN are in print only, as the data from the ten regional coordinating offices is not standardized, and in some cases no longer accessible online. In 1999, the NEWSPLAN Panel proposed a nationally-coordinated preservation program to identify and preserve "local newspapers in peril."¹³ The ten NEWSPLAN Regions working under the aegis of the British Library identified the 3,500 newspapers most at risk. The "NEWSPLAN2000" program received Heritage Lottery Funds to preserve up to 1,700 titles from the priority list. At the completion of the first phase of the effort (in operation from 2000-2005), 1,325 titles had been preserved, totaling more than 12.9 million pages transferred to microfilm. However, as noted by the NEWSPLAN Chair, "although many of the titles that were most at risk have been preserved by NEWSPLAN, a large number of titles will continue to be at risk until funding is secured for their preservation."¹⁴

Current legal deposit conditions in most European countries do not mandate the microfilming of all titles held at national and regional libraries. Where provisions do exist to ensure that incoming legal deposit content is microfilmed, those mandates were often put into place after the mid-20th Century and are typically forward facing, unable to accommodate the decades and even centuries worth of newspaper content already sitting in national collections. Like Sweden, Germany has had a concerted effort in place to microfilm all incoming newspaper titles since the 1970s, suggesting content published pre 1970 is, in large part, held only in print.¹⁵ Other national collecting institutions such as The British Library, whose legal deposit system has been in place since 1662, mandate that a copy of every publication be given to the Library by the publisher and still uses print as the default deposit format.¹⁶

Based on the data and anecdotal evidence assembled by CRL, it may be estimated that most (well above 50%) of European newspaper content published between 1669 and 1974 is held only in print.

¹⁰ National Library of Sweden, "Newspapers in Our Collections" <http://www.kb.se/english/collections/newspapers/>

¹¹ National Library of Sweden "Regina" <http://www.kb.se/english/find/catalogs/regina/>

¹² See general information and links at: <http://www.bl.uk/reshelp/bldept/news/newsplan/newsplan.html>

¹³ <http://www.bl.uk/aboutus/acrossuk/worknat/full/newsplan6.pdf>

¹⁴ <http://www.bl.uk/aboutus/acrossuk/worknat/full/newplanmin100628.pdf>

¹⁵ Europeana Newspapers "Featured partner: Hamburg State and University Library Carl von Ossietzky" <http://www.europeana-newspapers.eu/featured-partner-hamburg-state-and-university-library-carl-von-ossietzky/>

¹⁶ Center for Research Libraries *Framing a Common Agenda for Newspaper Digitization and Preservation: an ICON Summit Report and Outcomes* http://www.crl.edu/sites/default/files/event_materials/ICON_Summit-Report_and_Outcomes.pdf

Australia

In Australia, the National Library contributed significant resources to the preservation and digitization of national and regional newspaper titles under the Australian Newspaper Plan (ANPlan). ANPlan is a collaborative effort between the national, state and territory libraries to comprehensively collect, preserve, and provide access to Australian newspapers. A major resource that has been developed as a result of the ANPlan is the newspaper portal in NLA's online repository *Trove*, which provides access to 1,110 digitized historic newspaper titles from Australia.¹⁷

Microfilming is a key strategy of ANPlan. The intention of the project is to microfilm as many newspapers as possible before they become inaccessible. ANPlan libraries regularly report on the percentage of their hard copy newspaper holdings which have been microfilmed. It is estimated that a total of around 7,000 newspaper titles have been published in Australia.¹⁸ According to the ANPlan's 2014/15 annual report, approximately 71% (an average percentage of each territories' coverage) of all Australian newspaper content has been microfilmed (though the percentage varies by state/territory, ranging from 45% - 100%). Approximately 29% of newspaper holdings are still available only in print and remain at risk. ANPlan notes that many of these titles are currently held under conditions that do not meet the minimum qualifications set by the ANPLAN in terms of preservation and cataloging.¹⁹

Based on this very limited survey of the newspaper holdings of seven Western countries, an average 54% of newspaper materials are currently available exclusively in print format. The figures presented above suggest that a vast body of global newspaper material remains un-reformatted, and is at risk of becoming inaccessible as research and library communities integrate further into digital environments.

Strategic Action: improving transparency in newspaper collections

As demonstrated by the findings of this assessment, newspaper digitization has not yet significantly addressed or incorporated the preservation needs of library print collections. Original print newspapers that were not previously selected for microfilming by holding institutions continue to be neglected for preservation and access related transformations. The assessment's estimated average of 50% of library print holdings that have not been duplicated through reformatting indicates a vast body of newspaper content that remains scarcely available to library users. Print newspaper holdings are at risk of becoming permanently inaccessible and require the attention of the library community in order to prevent the loss of valuable primary source materials.

The world's print newspapers constitute an incomparable and irreplaceable source of critical historical and cultural evidence. Strategic action and a consensus on priorities for future digitization of newspaper collections by libraries and publishers is necessary to secure future access to these materials. In order to make strategic digitization and preservation related decisions possible for libraries, the community needs access to more transparent data on digitized collections, microfilmed materials and original print holdings

¹⁷ Trove <http://trove.nla.gov.au/newspaper/#>

¹⁸ National Library of Australia "REPATRIATION OF AUSTRALIAN NEWSPAPERS" <https://www.nla.gov.au/australian-newspaper-plan/repatriation-of-australian-newspapers>

¹⁹ National Library of Australia website, "ANPLAN Annual Report 2014/2015 Appendix E: Newspaper Titles Microfilmed" <https://www.nla.gov.au/anplan-annual-report-20142015-appendix-e-newspaper-titles-microfilmed>

sitting on library shelves and held by commercial providers. With access to granular data on newspaper collections and detailed information on the conditions, formats and issue level coverage of newspaper collections held by major aggregators, libraries will be able to independently assess the level of risk that exists in their own collections and can make informed choices on how to treat their content in digitization, microfilming and deaccessioning actions. CRL's ICON database offers a mechanism for sharing such information, and should be considered by libraries and commercial providers alike as a platform for exposing the outcomes of our shared effort.

Appendix A.

European Newspaper Repositories Surveyed for *The "State of the Art": A Comparative Analysis of Newspaper Digitization to Date*

- Austrian National Library (Austria)
- Belgium Bibliothèque royale de Belgique (Belgium)
- Danish State and University Library (Denmark)
- Kongelige Bibliotek (Denmark)
- British Library (United Kingdom)
- British Newspaper Archive, Ltd. (United Kingdom)
- King's College London (United Kingdom)
- National Library of Wales (Wales)
- National Library of Finland (Finland)
- Bibliothèque nationale de France (France)
- Bavarian State Library (Germany)
- Hamburg State Library (Germany)
- Landesbibliothek Wiesbaden (Germany)
- Staats- und Universitätsbibliothek Dresden (Germany)
- Staatsbibliothek zu Berlin (Germany)
- Thüringer Universitäts- und Landesbibliothek (Germany)
- Universitäts- und Landesbibliothek Bonn (Germany)
- Universitäts- und Landesbibliothek Darmstadt (Germany)
- Universitätsbibliothek Augsburg (Germany)
- Universitätsbibliothek Freiburg (Germany)
- Universitätsbibliothek Heidelberg (Germany)
- Koninklijke Bibliotheek (Netherlands)
- Isle of Man National Library (Isle of Man)
- Biblioteca comunale dell'Archiginnasio (Italy)
- Biblioteca Nazionale Braidense (Italy)
- Biblioteca Panizzi Reggio Emilia (Italy)
- Biblioteca Sportiva Nazionale (Italy)
- Biblioteca Universitaria di Pisa (Italy)
- Dr. Friedrich Teßmann Library (Italy)
- Norway National Library of Norway (Norway)
- National Library of Sweden (Sweden)
- Biblioteca cantonale di Lugano (Switzerland)
- Bibliothèque cantonale et universitaire – Lausanne (Switzerland)
- Swiss National Library (Switzerland)
- Société Neuchâteloise de Presse SA (Switzerland)

United States Newspaper Repositories Surveyed for *The "State of the Art": A Comparative Analysis of Newspaper Digitization to Date:*

- Library of Congress (Chronicling America)
- Library of Virginia
- Louisiana State University
- Montana Memory
- Ohio History Connection
- Portal to Texas History
- University of California-Riverside
- University of Hawaii
- University of Kentucky Libraries
- University of Oregon

- University of Utah
- Indiana State Library
- Connecticut State Library
- The State Historical Society of Missouri
- Washington State Library
- University of Illinois

Acknowledgments

James Simon, Vice President of Collections and Services, Center for Research Libraries

Bernie Reilly, President, Center for Research Libraries

Andrew W. Mellon Foundation

Brian Geiger, Director, CBSR, University of California, Riverside

British Library

Canadiana

Connecticut State Library

National Library of Austria

National Library of Wales

Indiana State Library

Ohio Digital Newspaper Program

Staats- und Universitätsbibliothek Hamburg

University of Kentucky Libraries

University of Oregon Libraries

References

Center for Research Libraries *The “State of the Art”: A Comparative Analysis of Newspaper Digitization to Date* (2015) from http://www.crl.edu/sites/default/files/d6/attachments/events/ICON_Report-State_of_Digitization_final.pdf

Europeana Newspapers *Europeana newspapers survey report* (2012) from <http://www.europeana-newspapers.eu/wp-content/uploads/2012/04/D4.1-Europeana-newspapers-survey-report.pdf>