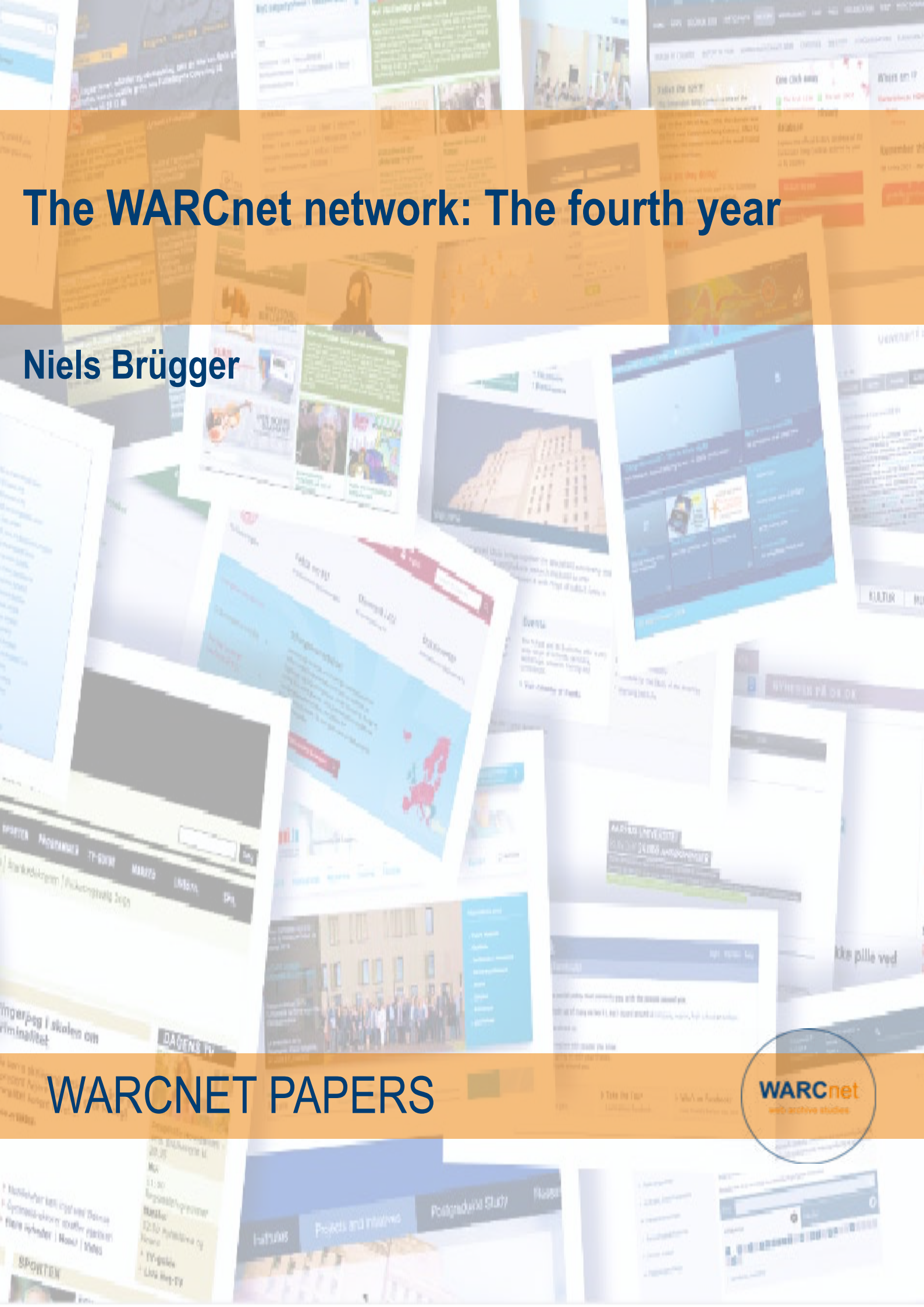
The WARCnet network: The fourth year

Niels Brügger



WARCNET PAPERS

WARCnet
web-archive studies

The WARCnet network: The fourth year

Niels Brügger

nb@cc.au.dk

WARCnet Papers
Aarhus, Denmark 2023

WARCnet Papers ISSN 2597-0615.

Niels Brügger: *The WARCnet network: The fourth year*
© The author, 2023

Published by the research network WARCnet, Aarhus, 2023.

Editors of WARCnet Papers: Niels Brügger, Jane Winters, Valérie Schafer, Kees Teszelszky, Peter Webster, Michael Kurzmeier.

Cover design: Julie Brøndum
ISBN: 978-87-94108-22-5

WARCnet
Department of Media and Journalism Studies
School of Communication and Culture
Aarhus University
Helsingforsgade 14
8200 Aarhus N
Denmark
warcnet.eu

The WARCnet network is funded by the Independent Research Fund Denmark | Humanities (grant no 9055-00005B).

DANMARKS FRIE FORSKNINGSFOND
INDEPENDENT RESEARCH FUND DENMARK

WARCnet Papers

Niels Brügger: *Welcome to WARCnet* (May 2020)

Ian Milligan: *You shouldn't Need to be a Web Historian to Use Web Archives* (Aug 2020)

Valérie Schafer and Ben Els: *Exploring special web archive collections related to COVID-19: The case of the BnL* (Sep 2020)

Niels Brügger, Anders Klindt Myrvoll, Sabine Schostag, and Stephen Hunt: *Exploring special web archive collections related to COVID-19: The case Netarkivet* (Oct 2020)

Friedel Geeraert and Nicola Bingham: *Exploring special web archives collections related to COVID-19: The case of the UK Web Archive* (Nov 2020)

Friedel Geeraert and Barbara Signori: *Exploring special web archives collections related to COVID-19: The case of the Swiss National Library* (Nov 2020)

Matthew S. Weber: *Web Archives: A Critical Method for the Future of Digital Research* (Nov 2020)

Niels Brügger: *The WARCnet network: The first year* (Jan 2021)

Susan Aasman, Nicola Bingham, Niels Brügger, Karin de Wild, Sophie Gebeil and Valérie Schafer: *Chicken and Egg: Reporting from a Datathon Exploring Datasets of the COVID-19 Special Collections* (Dec 2021)

Olga Holownia, Friedel Geeraert, Abbie Grotke, Jennifer Harbster and Gulnar Nagashybayeva: *Exploring special web archives collections related to COVID-19: The case of the Library of Congress* (Feb 2022)

Niels Brügger: *The WARCnet network: The second year* (Dec 2022)

Michael Kurzmeier: *Using a national web archive for the study of web defacements? A case-study approach* (Aug 2023)

Helle Strandgaard Jensen: *Any Teletubbies Caught in the Web?* (Aug 2023)

Niels Brügger: *The WARCnet network: The third year* (Aug 2023)

Kees Teszelszky: *Understanding the history of national domain crawls: mapping and archiving the national web domain in Denmark and France* (Aug 2023)

Niels Brügger: *The WARCnet network: The fourth year* (Aug 2023)

Valérie Schafer, Jérôme Thièvre and Boris Blanckemane: *Exploring special web archives collections related to COVID-19: The case of INA* (Aug 2020)

Niels Brügger, Valérie Schafer, Jane Winters (Eds.): *Perspectives on web archive studies: Taking stock, new ideas, next steps* (Sep 2020)

Friedel Geeraert and Márton Németh: *Exploring special web archives collections related to COVID-19: The case of the National Széchényi Library in Hungary* (Oct 2020)

Friedel Geeraert and Nicola Bingham: *Exploring special web archives collections related to COVID-19: The case of the IIPC Collaborative collection* (Nov 2020)

Caroline Nyvang and Kristjana Mjöll Jónsdóttir Hjörvar: *Exploring special web archives collections related to COVID-19: The Case of the Icelandic web archive* (Nov 2020)

Sophie Gebeil, Valérie Schafer, David Benoist, Alexandre Faye and Pascal Tanesie: *Exploring special web archive collections related to COVID-19: The case of the French National Library (BnF)* (Dec 2020)

Karin de Wild, Ismini Kyritsis, Kees Teszelszky and Peter de Bode: *Exploring special web archive collections related to COVID-19: The Dutch Web archive (KB)* (Nov 2021)

Michael Kurzmeier, Joanna Finegan and Maria Ryan: *Exploring special web archives collections related to COVID-19: The National Library of Ireland* (Feb 2022)

Emily Maemura: *Towards an Infrastructural Description of Archived Web Data* (May 2022)

Olga Holownia, Friedel Geeraert and Paul Koerbin: *Exploring special web archives collections related to COVID-19: The case of the National Library of Australia* (Dec 2022)

Helena Byrne, Beatrice Cannelli, Carmen Noguera, Michael Kurzmeier, Karin de Wild: *Looking ahead: after web (archives)?* (Aug 2023)

Friedel Geeraert, Marie Haškovcová, Luboš Svoboda and Markéta Hrdličková: *Exploring special web archives collections related to COVID-19: The case of the National Library of the Czech Republic* (Aug 2023)

Niels Brügger: *Why onsite meetings are important: Reporting from five Short-Term Network Stays* (Aug 2023)

Beatrice Cannelli: *Reporting from two Short-term Network Stays at The Institut National de l'Audiovisuel and the Bibliothèque nationale de France, and The Royal Danish Library* (Aug 2023)

WARCnet Special Reports

Sharon Healy, Helena Byrne, Katharina Schmid, Nicola Bingham, Olga Holownia, Michael Kurzmeier, Robert Jansma: *Skills, Tools, and Knowledge Ecologies in Web Archive Research* (Aug 2022)

Sharon Healy, Helena Byrne, Katharina Schmid, Juan-José Boté-Vericad, Lanna Floody: *Towards a Glossary for Web Archive Research: Version 1.0* (Aug 2023)

Sharon Healy, Helena Byrne: *Scholarly Use of Web Archives Across Ireland: The Past, Present & Future(s)* (Aug 2023)

All WARCnet Papers and WARCnet Special Reports can be downloaded for free from the project website warcnet.eu.

The WARCnet network: The fourth year

Niels Brügger

Abstract: This WARCnet Paper provides an overview of the WARCnet network's activities in the first eight months of 2023, the fourth and last year of the project period.

Keywords: web, WARCnet, activities, 2023

TRAVELLING AND NETWORKING ONSITE

The WARCnet network was scheduled to run for two years, 2020-2021. However, both years were marked by the outbreak of the COVID-19 pandemic, and therefore the project was launched in Zoom, and for one year and a half project activities did not get out of the shadows of COVID-19. In many ways the pandemic was a challenging period, and it obviously affected a networking project like WARCnet in a negative manner, but there were also a few upsides of it, in particular that the project period could be exended because no funding could be spent on onsite meetings — in total four extensions were granted by the funder, the Danish Independent Research Fund Denmark | Humanities, and therefore activities could also take place in the first eight months of 2023, including one network meeting onsite and the allocation of several Short-Term Network Stays (STNS).

The following paragraphs outline the network activities that have taken place in the first eight months of 2023.

NETWORK MEETING AND AUTHOR GROUPS

In 2023 no members left the network, and therefore WARCnet had the same cohort of participants as the previous year: 37 participants, including 28 researchers, 7 web archives, 1 IT developer, and a transnational international web archiving organisation, the International Internet Preservation Consortium (IIPC). Following the shift in the organisation of the network that was started in 2022 the focus of network activities in 2023 was exclusively on 'Author Groups', that is the groups that were formed depending on which chapter(s) participants would contribute to the planned *Routledge Companion to*

Transnational Web Archive Studies that constitutes an important tangible outcome of the network activities. In practice this meant that the Working Groups (WGs) that were previously one of the primary organisational entities only played a minor role in 2023.

In 2023 only one network meeting was organised, and it was exclusively about the chapters to be included in the *Routledge Companion to Transnational Web Archive Studies*. The meeting took place in Marseille 7 June (onsite only), back-to-back with the fifth RESAW Conference 'Exploring the Archived Web During a Highly Transformative Age' (5/6 June) which all meeting participants attended. RESAW, a Research Infrastructure for the Study of Archived Web Materials, is a leading international community established in 2012, aiming at promoting a collaborative European research infrastructure for the study of archived web materials, and it organises biennial conferences in Europe (see resaw.eu). In relation to the RESAW Conference WARCnet sponsored a small reception.

The WARCnet meeting was organised by Aarhus University, and 13 chapter authors participated, including network members as well as non-members since all keynotes from the project period had also been invited to contribute a chapter.

Before the meeting all participatants had submitted a first draft of their chapter(s), and the aim of the meeting was to give peer feedback to these drafts. During this feedback session the following format was used:

- all chapters were sent to all participants,
- each participant was asked to read, present and comment on one of the other author's chapter,
- presenting and commenting could include adressing questions such as: Is the main topic clear? Are the data relevant and well-described? Are the methods relevant? What are the two main insights and the two main things to improve?,
- then followed comments from the author, and an open discussion with all participants,
- each feedback session was rounded off with a summarising of the main take-away points from the conversation,
- commenters could make a few slides in case they wanted to.

Also, a minutes keeper was hired in to provide extensive minutes that were communicated to all participants after the meeting.

After the meeting chapters that were discussed would go through the single blind review procedure as all other chapters. Finally, the procedure for the four 'Conversations' that are to be placed between the book's five sections was discussed. A 'Conversation' takes up one or more topics or discussions addressed in the preceding section and/or point forward to issues being discussed in the following section, and it is to be made in a shared online document where chapter authors are invited to discuss questions formulated by the editors and/or suggested by the authors themselves.

THE WARCNET BOOK

The writing and editing of the *Routledge Companion to Transnational Web Archive Studies* has been one of the main network activities in 2023. The book is edited by Susan Aasman, Anat Ben-David, and Niels Brügger, it will be published by the international publisher Taylor & Francis/Routledge in 2024, and it is estimated to be 430 pages long.

The book comprises a Foreword, an Introduction by the editors, 23 chapters, and four 'Conversations', and it is divided in five sections: (1) entire national web domains from a transnational perspective, (2) the COVID-19 crisis as a transnational event, (3) methods and skills in web archive studies, (4) politics of web archives as collections, and (5) what to do with web archives? At the time of writing 19 chapters by network members, keynotes and one network-external contributor have been received and have gone through a process of internal single blind review, and the last 4 chapters are expected to be submitted mid-September 2023.

SHORT-TERM NETWORK STAYS

In 2022 the lifting of travel limitations imposed by the pandemic had made it possible to use WARCnet's Short-Term Network Stays where a researcher, a web archivist or IT staff by one of the participating research institutions and/or web archives could visit another network member's institution. In 2023 the Short-Term Network Stay Committee allocated four network stays, two in the UK, one in Luxembourg, and one in the Netherlands, all granted to network members from Denmark.

DISSEMINATION AND OUTREACH

In 2023 the WARCnet network has continued disseminating knowledge related to network activities. Communication has taken place on the project website at warcnet.eu, and on social media platforms, and results have been disseminated in the network's two publication series: the WARCnet Papers, and the WARCnet Special Reports. Also, several network initiated conference presentations have been submitted and accepted, including to the RESAW Conference mentioned above.

The project website will remain online (not updated) for two years after the project has ended. To preserve WARCnet's activities for the future all content on the project website, including all project publications, will be saved in the US based Internet Archive (archive.org) as well as in the Portuguese web archive Arquivo.pt (arquivo.pt) where it will be freely available. Also, a copy will be downloaded by the UK Web Archive at the British Library where it will be made available in the Web Archives and Digital Preservation subsection of the IT Collection (https://www.webarchive.org.uk/en/ukwa/collection/520). Thus, there should be several opportunities to find material from the project in the future.

warcnet.eu, and social media

The website warcnet.eu (hosted by Aarhus University) distributes information about the project. Since the WARCnet project has only run for eight months in 2023, all the statistics below cannot be compared directly. Therefore, where relevant comparisons are made using average numbers or percentages.

In the first eight months of 2023 warcnet.eu had 513 visitors (64 per month) which is a considerable average decrease compared to 2022 (4,765 visitors, 397 per month). Visitors still come from a variety of countries, but when comparing the visitors' country of origin from 2022 to 2023 (Figure 1) it is striking that the number of visitors from the United States has been halved and more visitors come from France, whereas for the rest the variations are minor.

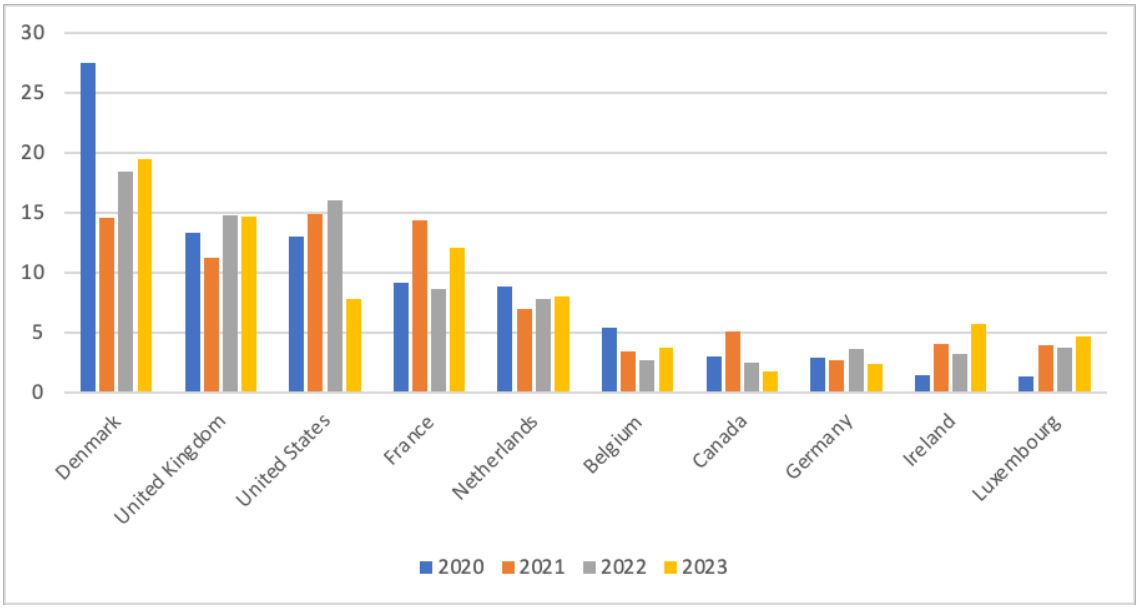


Figure 1: Visitors on warcnet.eu, 2020-2023, countries (% of total).

Compared to 2022, in 2023 the number of users accessing the website from a stationary computer has increased from 88% to 94%, and only 6% used either a mobile phone or a tablet to access the website which is the lowest figure in the entire project period (15% in 2020, 9% in 2021, 12% in 2022). Not surprisingly the landing page is still the most popular, but in contrast to 2022 where the second most popular web page was the page with information about the Aarhus Closing Conference 2022, followed by the web page about WARCnet, and the WARCnet Team page, in 2023 the web page with WARCnet Papers and Special Reports is second on the list, and it is almost as popular as the front page. This could suggest that the many WARCnet Papers and Special Reports published towards the end of the project period in August 2023 has generated significant trafic on the website. Third, fourth and fifth on the list for 2023 are almost similar to what was the case in 2022: the web page about WARCnet, the WARCnet Team page, and the web page about Working Groups. This distribution of most popular web pages is mirrored in which web pages are the most used entry points to the website, except that the web page

about the 'WARCnet Twitter Challenge' is fifth on the list in 2023, although the Twitter Challenge ended in October 2022.

In all previous years most users typed in the web address warcnet.eu in their browser to reach the website whereas search engines were not much used, although their use has increased during the project period. However, this has dramatically changed in 2023 (Figure 2) where less than 30% seem to know the web address warcnet.eu beforehand, which is probably also the reason why the use of search engines has almost doubled compared to 2022. The number of visitors coming from social media is still quite small (and as in previous years the vast mojority come from Twitter), but visitors being refered to warcnet.eu via a link on another web page has tripled. This overall picture indicates that in the first project years visitors knew where to find WARCnet's website, whereas towards the end of the project period the web address is less familiar to visitors.

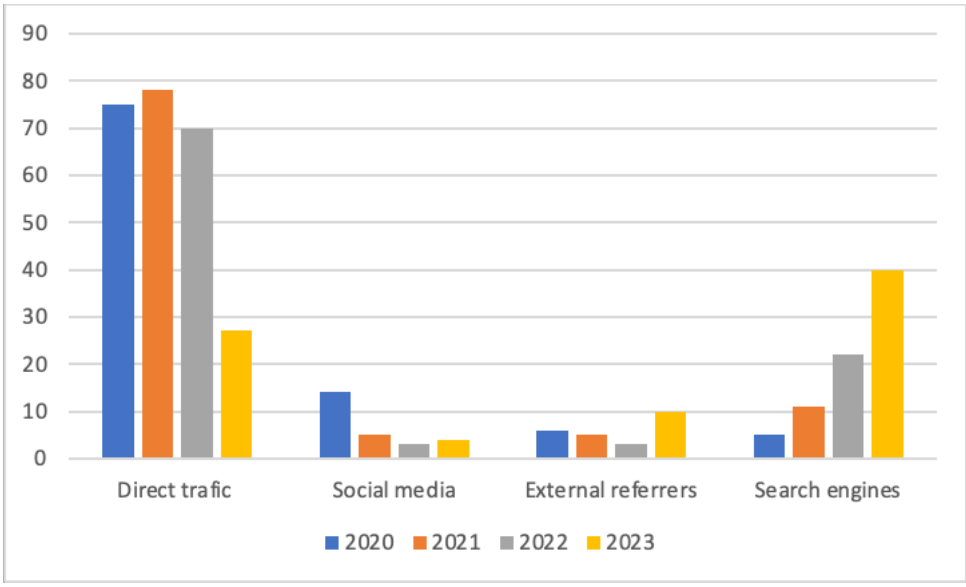


Figure 2: Visitors come to warcnet.eu from various sources, 2020-2023 (% of total) (figures for 2023 do not add up to 100% because a new category 'internal referrers' was only introduced in 2023).

Finally, when looking at the distribution of visitors over the year (Figure 3) visitors tend to be evenly distributed during the first seven months of 2023, with a small decrease in February which is nicely in line with the fact that only one internal network meeting has been held in this period. However, August stands out with an increased number of visitors, probably because a large number of WARCnet Papers and Special Reports were published in this month, as also mirrored in the above mentioned high popularity of the webpage where these publications can be found.

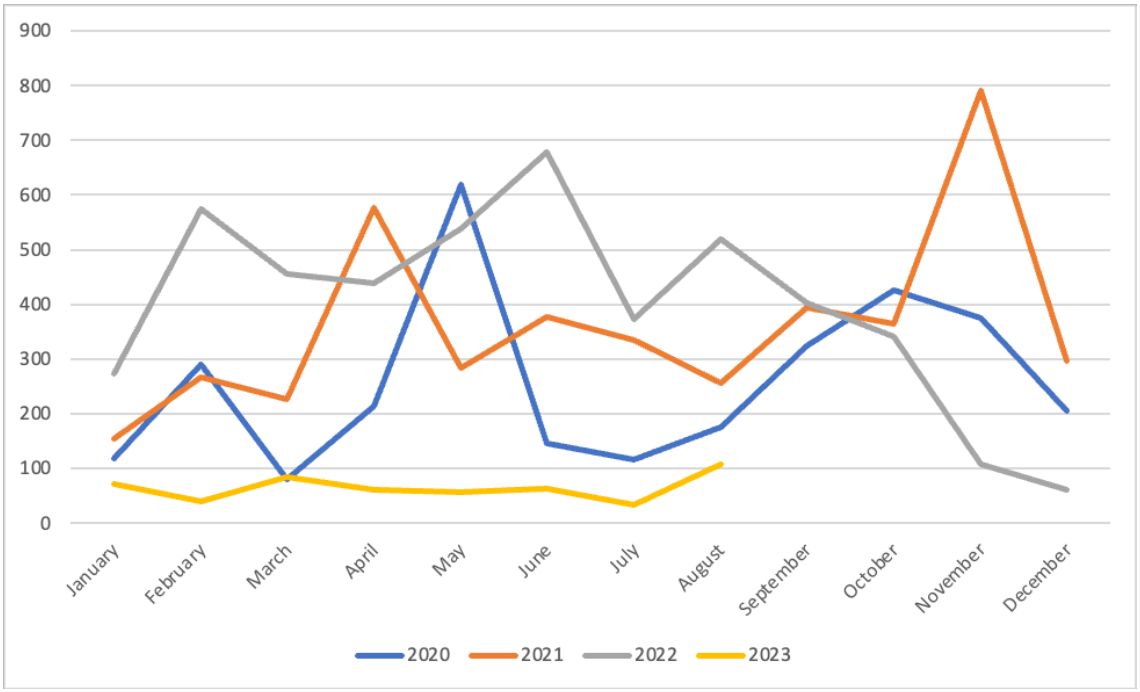


Figure 3: Visitors on warcnet.eu per month, 2020-2023 (number of visitors).

WARCnet also has a presence on Facebook, Twitter, YouTube, and Slideshare, and in 2023 the number of followers has roughly remained the same as in 2022.

WARCnet Papers, and WARCnet Special Reports

The WARCnet Papers disseminate results and other network related material. In 2023 nine papers were published, all in August, the last month of the project period, including (a) a co-authored paper with the contributions to a keynote panel with WARCnet's early career scholars, (b) two papers that provide overviews of the third and fourth years' network activities, and (c) a number of publications reporting from one or more STNS.

The WARCnet Special Reports is a publication series that was initiated in 2022 with a view to offering a publication outlet for longer reports and similar, and in 2023 two WARCnet Special Reports have been published, including a more than 200 pages long report about the scholarly use of web archives across Ireland, and a report developing a glossary of terms and concepts for web archive research (67 pages).

Download figures for 2023 are influenced by the fact that all publications from 2023 were published in the last month of the project period. In 2023 all WARCnet Papers were visited 148 times, of which 113 were visits to papers that had been published in previous years and 35 were to publications only published in August 2023 which is a relatively high download rate for less than one month compared to previously published papers' download rate in 2023. As indicated in Figure 4 there is still a steady interest in the topics covered by WARCnet Papers since all previously published papers continue to be visited.

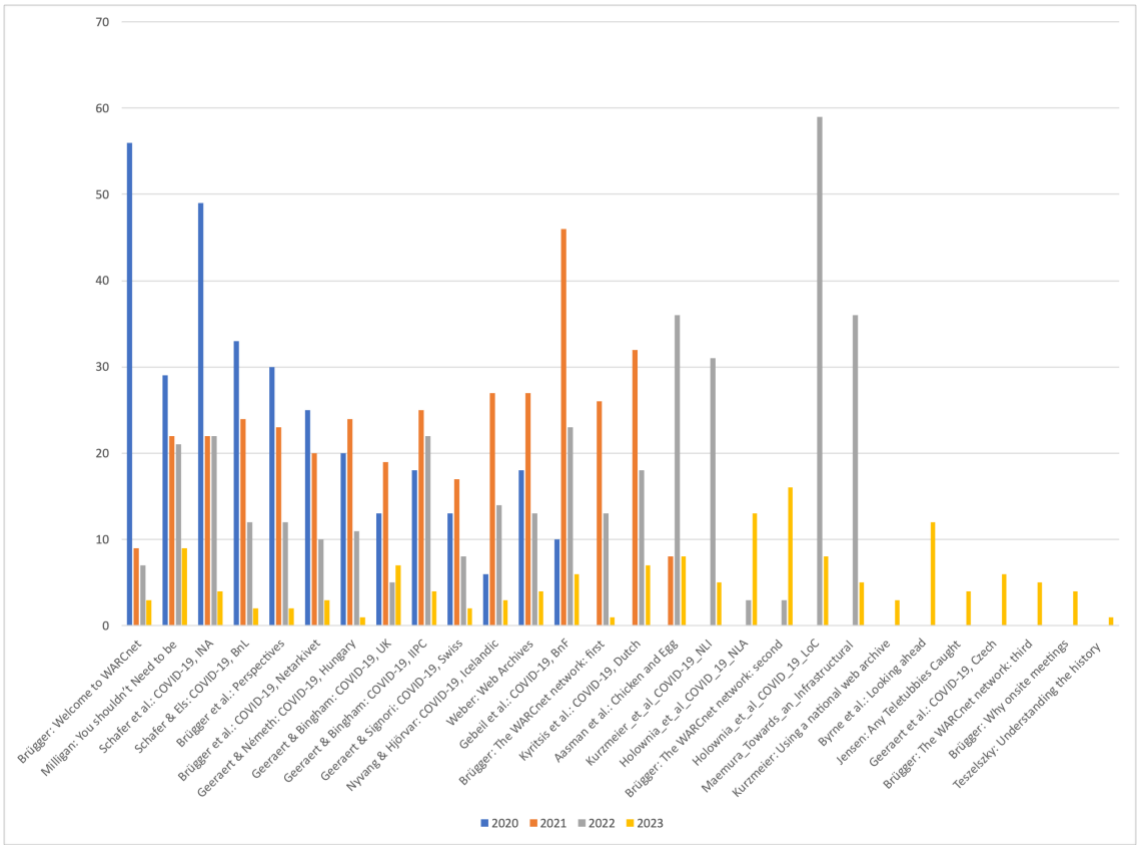


Figure 4: WARCnet Papers, number of visits on warcnet.eu per year, 2020-2023 (publications in chronological order, from left to right).

In total, from 2020 to 2023 WARCnet Papers have been visited 1,218 times, and the most popular paper has been visited almost 100 times, as shown in the top-20 in Figure 5.

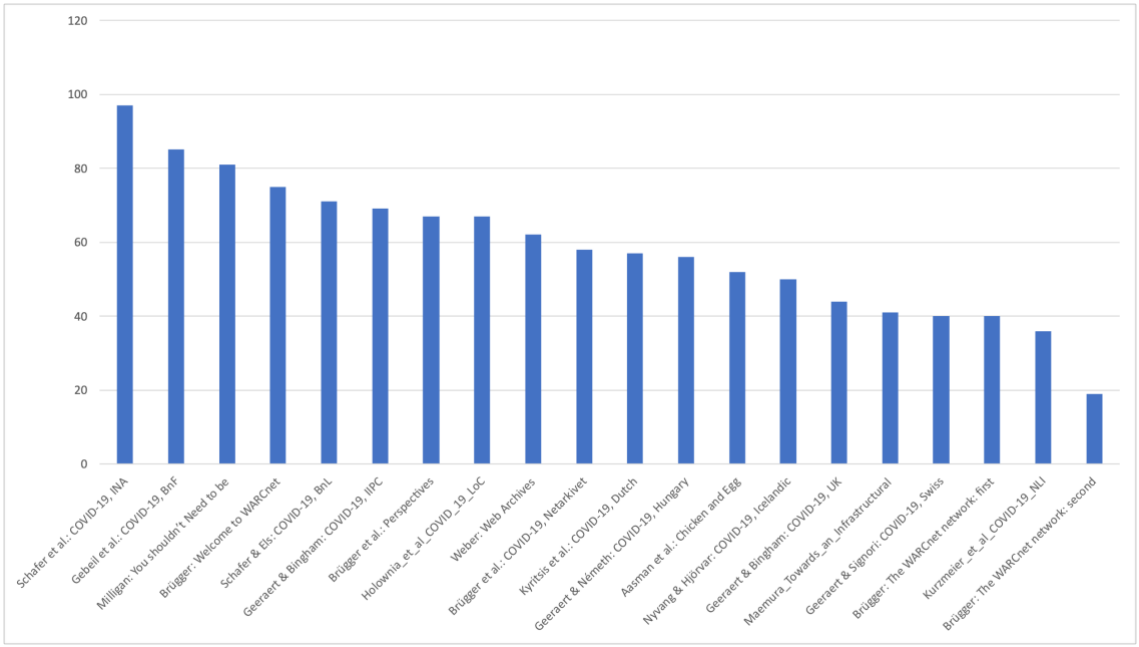


Figure 5: WARCnet Papers, top-20 visited papers 2020-2023.

As for WARCnet Special Reports they have been visited 33 times in 2023, 20 times for the report published in 2022, 13 times for the two reports published in August 2023. In total all three Special Reports have been visited 91 times.

MANAGEMENT

WARCnet's Steering Group oversees the management of the WARCnet network, and it has met once in 2023. In continuation of the diminishing importance of the WGs WARCnet's Coordination Group that should ensure the coordination between the WGs has not met in 2023. The Short-Term Network Stay Committee that allocates STNS has evaluated applications on a rolling basis. Finally, the administration and management of the project is supported by two student helpers, both from Aarhus University.

WARCNET IN THE FUTURE

The WARCnet network has now reached the end, with a number of significant results, among others a lively, inclusive and highly productive international network of academics and web archivists working with web archives, a large number of network meetings documented through presentations on warcnet.eu, a well-attended closing conference, more than 30 project publications in the form of WARCnet Papers and Special Reports, and an edited book of more than 430 pages to be published in 2024. Hopefully, all these results will constiute important and valuable stepping stones for future research and web archive activities within this field of research. At least, the foundations of making this happen have been formed.

The network started in the shadows of COVID-19, which made it difficult to establish a network that was conceptualised as a network based on physical onsite meetings. However, despite being locked up in Zoom for almost two years with all the well-known downsides network members made a great effort to make the network work, and it is still surprising that so many activities could be made during these unfamiliar circumstances. This would never have happened without the positive approach of all participants. Also, it is worth highlighting that COVID-19 also came with a number of opportunities for improving the network, in particular extending the project period to almost the double and increasing the number of network members beyond the initial group by doubling the number of participants that started out. Regarding the many new network members it is particularly beneficial for the research field that the network was able to provide an academic community for a large number of early career scholars.

Although the WARCnet network finaly did get out of the shadows of COVID-19 the activities of the project have been marked by the pandemic in negative as well as positive ways, and one can only speculate how the network would have developed had it only been for two years and limited to the initial network members. We will never know.

WARCnet Papers is a series of papers related to the activities of the WARCnet network. WARCnet Papers publishes keynotes, interviews, round table discussions, presentations, extended minutes, reports, white papers, status reports, and similar. To ensure the relevance of the publications, WARCnet Papers strives to publish with a rapid turnover. The WARCnet Papers series is edited by Niels Brügger, Jane Winters, Valérie Schafer, Kees Teszelszky, Peter Webster and Michael Kurzmeier. In cases where a WARCnet Paper has gone through a process of single blind review, this is mentioned in the individual publication.

The aim of the WARCnet network is to promote high-quality national and transnational research that will help us to understand the history of (trans)national web domains and of transnational events on the web, drawing on the increasingly important digital cultural heritage held in national web archives. The network activities run in 2020-22, hosted by the School of Communication and Culture at Aarhus University, and are funded by the Independent Research Fund Denmark | Humanities (grant no 9055-00005B).

WARCNET PAPERS

INDEPENDENT RESEARCH FUND DENMARK

warcnet.eu warcnet@cc.au.dk twitter: @WARC_net facebook: WARCnet
youtube: WARCnet Web Archive Studies slideshare: WARCnetWebArchiveStu