The WARCnet network: The third year

Niels Brügger

WARCnet
web-archive studies

The WARCnet network: The third year

Niels Brügger

nb@cc.au.dk

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)

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)

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* (August 2022)

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

The WARCnet network: The third year

Niels Brügger

Abstract: This WARCnet Paper provides an overview of the WARCnet network's many activities in 2022, the third year of the project period.

Keywords: web, WARCnet, activities, 2022

NETWORKING OUTSIDE THE SHADOW OF COVID-19

The WARCnet network started shortly after the outbreak of the COVID-19 pandemic in 2020. Thus, the first year and a half of the project took place in the shadows of COVID-19, and it was not until November 2021 that the fourth network meeting could be held as a hybrid session, with 13 network members meeting onsite in Aarhus and the rest of the group participating via Zoom.

2022, the third year of the project, finally allowed networking outside the shadow of COVID-19: network meetings onsite could now take place, one of the funding instruments of WARCnet, the so-called Short-Term Network Stays, could now be used, and the closing conference could be held as planned. This was only made possible because of a third extension of the project period by the funder, the Danish Independent Research Fund Denmark | Humanities. This extension allowed WARCnet to continue until 30 June 2023. However, towards the end of 2022 it became clear that some activities would have to be scheduled after this date, and therefore a short fourth (and last) extension was granted until 30 August 2023. Thus, the project is ending more than one and a half year later than initially planned.

The following paragraphs outline the network activities that have taken place in 2022, with a focus on four major activities: an onsite meeting, the closing conference, the 'WARCnet book', one of the tangible results of the project, and Short-Term Network Stays.

NETWORK MEETINGS, WORKING GROUPS, AND AUTHOR GROUPS

In 2022 WARCnet had 37 participants, including 28 researchers, 7 web archives, 1 IT developer, and a transnational international web archiving organisation, the International

Internet Preservation Consortium (IIPC). In 2022 one member has left the network. In previous years the primary organisational entities of WARCnet have been network meetings and Working Groups (WGs). However, in 2022 the focus gradually shifted from working in WGs to preparing chapters for the planned 'WARCnet book', and since this work did cut across WGs the importance of the WGs diminished (see below about the preparation of the planned book).

Network meetings

In 2022 two network meetings were organised, one 13-15 June in London (primarily onsite, but also with online participation), and a very brief meeting 16 October in Aarhus, right before the closing conference. The two meetings were organised by the University of London, School of Advanced Studies, and by Aarhus University, the School of Communication and Culture, Department of Media and Journalism Studies, respectively.

The June meeting took place under the heading 'Web archive features and futures', and the aim was to: (a) consider the key features of web archives and web archive studies that have been highlighted by the WARCnet project, (b) explore how a greater understanding of these features will help to inform future directions for web archives. There were six types of presentations: (a) keynotes that were open for participants outside the network, (b) WG sessions, (c) WG presentations to other WGs, (d) a Tools Bazaar, where participants could share and demonstrate tools, scripts, datasets, and workflows they are using, (e) sessions about the future of and after the WARCnet network, (f) a public discussion session on the resilience of the born digital. All in all 12 presentations were made by network members, 22 network members participated onsite, 11 took part in the meeting via Zoom.

The meeting included two keynotes, "From Teddy Bears and Cyborg Witches to Oysters and Pink Elephants: A xenofeminist remedy for the mal d'archive and other binary troubles" by Saskia Huc-Hepher, University of Westminster, and "Humans are ready to use some web archives": Working with the Wayback Machine and Wikipedia's history pages in the not-so-digital humanities" by Sandra Folie, University of Vienna.

The Tools Bazar included five 'stations' between which participants could circulate. 'Stations' were manned by network members or people from outside the network who introduced different software tools, including 'Wayback Machine scraper' (Caio Mello), 'Twitter analysis dashboard' (Kunika Kono, Martin Steer), 'Archived Unleashed Toolkit' (Frédéric Clavert), and 'Warc2Corpus' (Sebastian Gassner). Also, a new round of the Twitter Challenge was initiated, under the heading 'ALIENS on the web (or in the web archive)'.

The brief meeting 16 October in Aarhus was mainly a business meeting to discuss how to plan the work during the last months of the project. Also, the meeting included a brief status about and a discussion of the following steps for the planned 'WARCnet book'.

Working Groups, and Author Groups

Each network participant could join one (or more) of the six existing Working Groups. All WGs, including their aims and tasks, are described in detail on the network's website at https://cc.au.dk/en/warcnet/working-groups. WGs continued to be one of the main structural entities in the project at the above mentioned meeting in London in June, but for the rest of the year (and the rest of the project period) network members formed 'Author Groups', depending on which chapter(s) they would like to contribute to the planned 'WARCnet book'. WGs continue to exist, but much of their work is now taking place in the more focused 'Author Groups' that in some cases also cut across WGs. Thus, 'Author Groups' constitute an overlay to the WGs, and in practice, in the last part of 2022, 'Author Groups' have become the primary organising unit of the project.

CLOSING CONFERENCE

Despite the fact that the project would not be terminated until mid-2023 WARCnet organised a closing conference to take place in Aarhus, Denmark, 17-18 October 2022. The dates for the closing conference were already scheduled late 2021, before the third extension of the project period, after having been postponed already once, and in order not to make advertising of the conference even more confusing it was decided to keep the planning of a closing conference several months before the actual project closing. Another strong argument to keep the initial dates for the closing conference was that it would take place during the two days leading up to the conference of the European Communication Research and Education Association (ECREA), which also took place in Aarhus. Thus, the closing conference could function as a lead in to the ECREA conference which could potentially increase the number of participants.

An Organising Group was set up to draft the Call for Papers, to run the review process of abstracts, to draft the conference program, and to take care of the planning on the ground. And a Scientific Committee with nine international academics was established with a view to overseeing the scientific quality of the submissions. The Call for Papers was posted widely already by the end of 2021. It was decided that the conference should be onsite only, and that it should be open for contributions from researchers, web archivists, IT-developers, and practitioners from inside and outside the network who could contribute to the field of study of WARCnet. Submissions could focus on web archives and at least one of the following fields: (a) studies of web domains, (b) studies of transnational events with web archives, (c) studies of events on the Web, (d) digital methods and tools, (e) research Data Management, (f) data formats for web archive studies, (g) web archiving practices and challenges. Submissions were welcomed from all disciplines within the humanities and social sciences, including media and communication, history and digital humanities, and data scientists, web archivists, web curators, and practitioners were welcomed too.

The heading for the closing conference was 'Web archive studies researching web domains and events', 24 submissions were received, including individual papers as well

as panels, and 18 papers and four panels were accepted. The conference was attended by 59 participants, mainly from European countries, but also some from outside of Europe.

The conference included one keynote and a 'keynote panel'. The keynote was "From 'History Web' and 'Web History' to the 'History of the Web'? Reflections from a project studying the origins of the World Wide Web" by Gabriele Balbi, Università della Svizzera italiana, Switzerland, and in the keynote panel "Looking ahead: after web (archives)?" five early career scholars from the network were invited to address at least two of the following five questions: (a) what do you consider the greatest challenges to having web archive studies advance?, (b) how to develop public and societal engagement with web archives?, (c) is AI the future of web archiving and web archives studies?, (d) how can web archives be used alongside other sources, both analogue and digital?, (e) what new skills do web archivists and researchers of web archives of the future need (which they don't possess now)? Participants in the 'keynote panel' were Helena Byrne (the British Library), Beatrice Cannelli (School of Advanced Study, University of London), Carmen Noguera (Luxembourg Centre for Contemporary and Digital History, University of Luxembourg), Michael Kurzmeier (Department of Digital Humanities, University College Cork), Karin de Wild (Department of Art History, Leiden University).

Also, the final round of the Twitter Challenge was launched, under the heading 'The future of the web in the past and the past web in the future'. And since this was the final Twitter Challenge the winner of the winners from all five rounds was also found and was awarded 'the Great Raider of the Lost Web Award'.

Finally, the conference also included social events, including visits at the Royal Library and the Danish web archive Netarkivet, at AROS, the arts museum of Aarhus, and at a local gin destillery.

THE WARCNET BOOK

Already in the Spring of 2021 members first started to discuss finishing the project with an edited volume to be published at an international publisher. To push the process forward an Editorial Task Force was formed in May 2021 (Susan Aasman, Anat Ben-David, and Niels Brügger), and the task of this group was to initiate a debate if/how members should aim at publishing an edited volume as one of the outputs of the network, and in case it was decided to go for an edited volume the Task force would be the editors of this book.

During 2021 and 2022 the Editorial Task Force drafted a couple of scenarios as to the book's structure that were discussed at business meetings at the fourth and fifth network meetings. Based on these discussions a 35 pages book proposal for an edited volume entitled *Seeing the past and the present through web archives: A transnational perspective* was finalised by July 2022. The proposed book was 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? — and it included a Foreword, an Introduction by the editors, and 25 chapters. To have the book mirror the

network's activities the best possible all keynotes who had presented during the project period were also invited to contribute a chapter, thus, 7 of the chapters would be authored by keynotes. Also, one network external author was invited in to cover legal aspects of the researcher use of web archives. In addition to the chapters themselves it was decided to insert four 'Conversations' between the five sections. A 'Conversation' should take 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 a 'Conversation' is made in a shared online document where chapter authors are invited to discuss questions formulated by the editors.

The book proposal was submitted for review at the international publisher Taylor & Francis/Routledge, and based on four positive reviews the book proposal was revised, and a contract was signed in December 2022. In addition to various adjustments the title of the book was changed to *The Routledge Companion to Transnational Web Archive Studies*. The book is scheduled to be published in 2024, it is estimated to be 430 pages long, and the drafting of the book chapters constitute one of the major activities of the network until project end.

SHORT-TERM NETWORK STAYS

Travel limitations imposed by the pandemic had made it impossible to initiate one of the funding instruments of WARCnet, the so-called Short-Term Network Stays (STNS). STNS are intended to support research and knowledge exchange among the members of the network through short-term visits. An STNS involves the hosting of a researcher, a web archivist or IT staff by a participating research institution and/or web archive. When stepping out of the shadow of COVID-19 in 2022 it finally became possible to launch this planned instrument, and the Short-Term Network Stay Committee could therefore allocate six network stays, three in France, and three in Denmark, granted to network members from the UK, the Netherlands, Ireland, and Denmark.

DISSEMINATION AND OUTREACH

In 2022 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 (no Newsletters were published in 2022), whereas results have been disseminated in the network's publication series WARCnet Papers that in 2022 was supplemented by a new publication outlet: WARCnet Special Reports (see more below). Also, a number of network initiated conference presentations have been submitted and accepted. Finally, also in 2022 WARCnet's Twitter Challenge has constituted an example of outreach to a wider public.

warcnet.eu, and social media

On the website warcnet.eu (hosted by Aarhus University) one can find relevant information about the project's life, including presentations from all meetings. In 2022 warcnet.eu had 4,765 visitors which is an increase of approximately 10% compared to 2021 (4,324 visitors). Visitors still come from a variety of countries, but compared to the first year (2020) where a quarter of the visitors came from Denmark, in the following years this was evened out since the number of visitors from the United Kingdom and the United States has become much closer to that of Danish visitors (now decreased), and partly also to the number from France. As for the other countries in the top-10 of visitors figures remain relatively stable over the years.
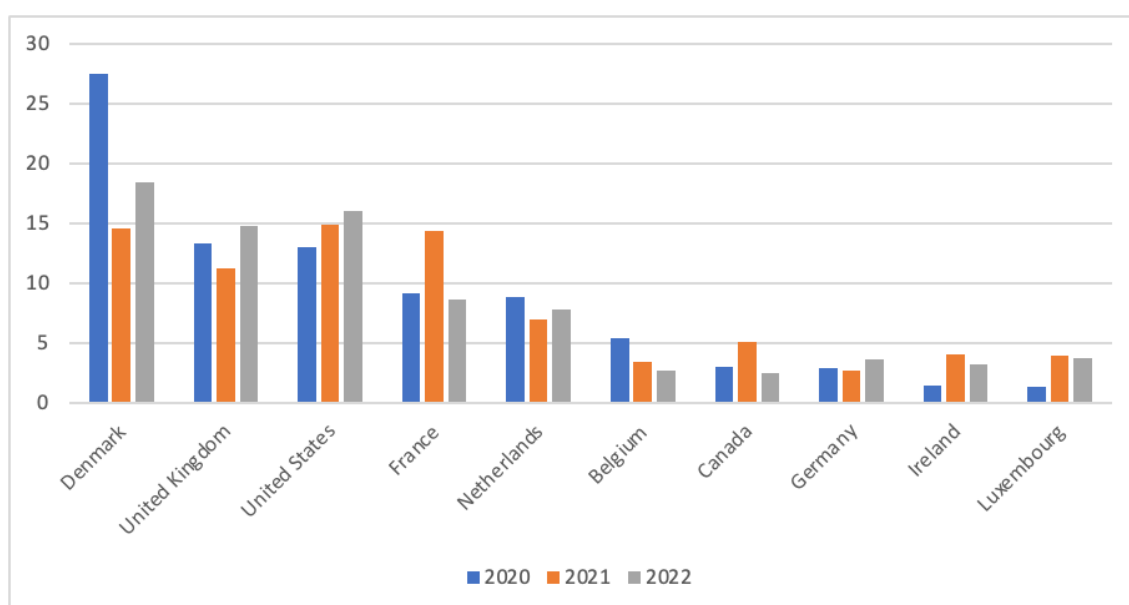
Figure 1: Visitors on warcnet.eu, 2020-2022, countries (% of total).

Compared to 2021, in 2022 the number of users accessing the website from a stationary computer almost remained the same (88% compared to 91%), and only 12% (9% in 2021) used either a mobile phone or a tablet. Not surprisingly the landing page is the most popular, and then follows the web page with information about the Aarhus Closing Conference 2022, the web page about WARCnet, the WARCnet Team page, and the web page related to the meeting in London June 2022. Surprisingly, the WARCnet Papers web page is only visited by 2.9% of the visitors, in contrast to previous years where it was the second most popular web page. This overall picture of most popular web pages is also mirrored in which web pages are the most used entry points to the website.

   In 2020 and 2021 most users simply typed in the web address warcnet.eu in their browser to reach the website and this is still the case in 2022. However, it is worth noting that in 2022 compared to previous years fewer visitors come from social media, whereas more visitors are referred to the website from a search engine. In fact, the tendency of social media and search engines swapping place between 2020 and 2021 is continued,

as can be seen in Figure 2. Among those coming from social media almost all of them still came from Twitter, which is similar to 2020 and 2021.
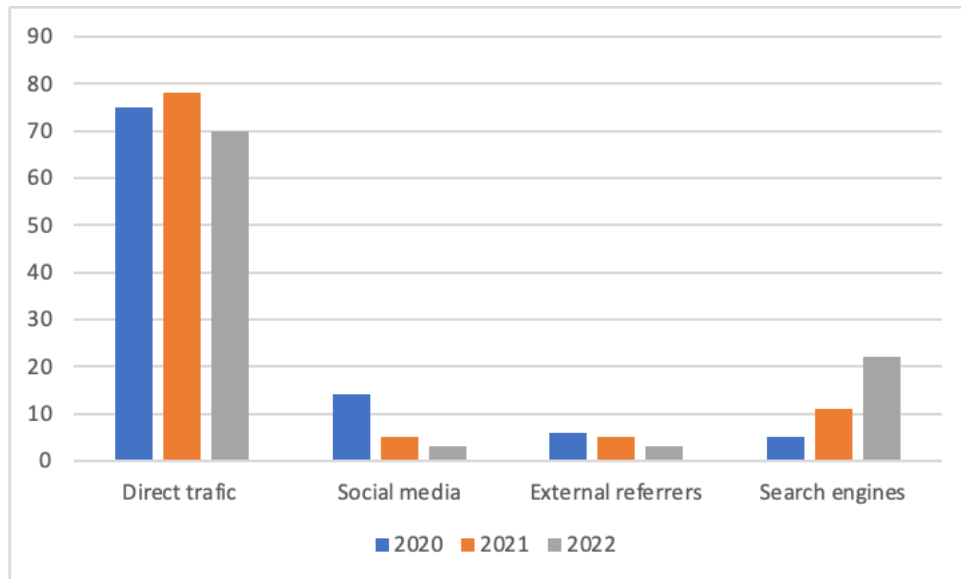

Figure 3: Visitors come to warcnet.eu from various sources, 2020-2022 (% of total).

Finally, when looking at the distribution of visitors over the year (Figure 3) in 2020 as well as in 2021 the main visitor activity was related to network meetings — May and November in 2020, and April and November in 2021 — but this is only partly the case in 2022: many users visit the website before the meeting in June in London, but surprisingly few visit the website in the period leading up to the closing conference in October, although there is a small peak in August. This is even more surprising since the web page with information about the closing conference is one of the most visited web pages.
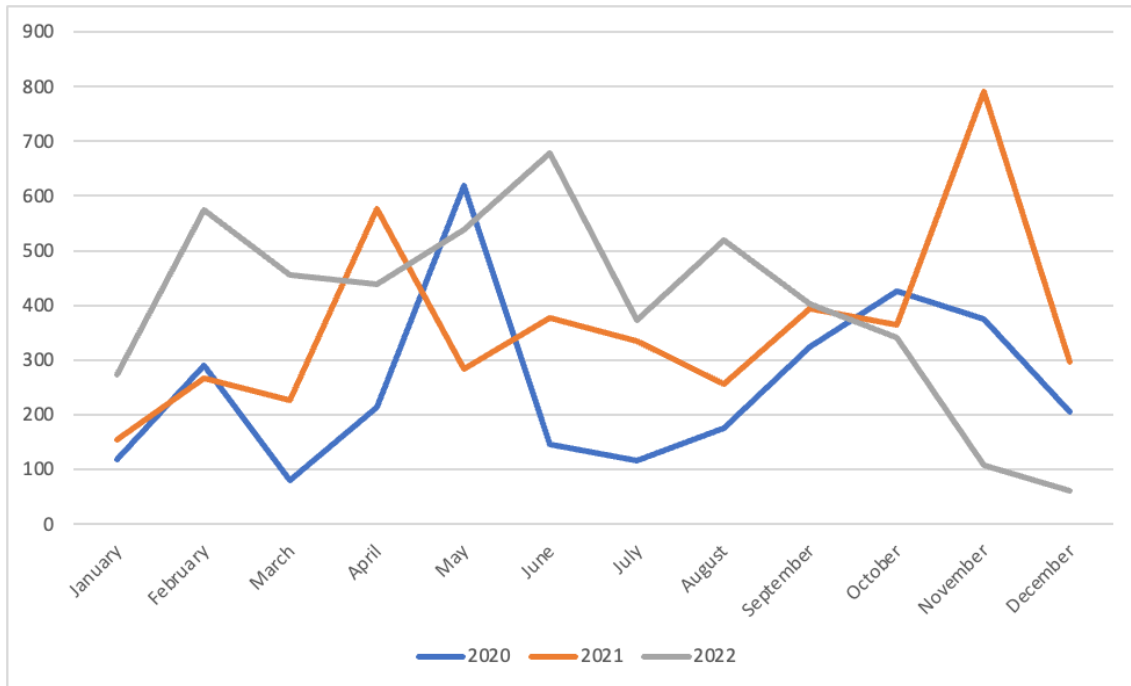
Figure 3: Visitors on warcnet.eu per month, 2020-2022 (number of visitors).

WARCnet has a presence on Facebook, Twitter, YouTube, and Slideshare. The Facebook page still has 15 followers, whereas the Twitter profile has increased its number of followers from 278 to 347. The YouTube chanel has also increased its number of subscribers, from 9 to 22 and has been viewed 224 times, which is a significant decrease compared to the 511 views in 2021. Figures for Slideshare are not available for 2022.

WARCnet Papers, and WARCnet Special Reports

WARCnet Papers is the project's publication series where network related material that does not fit into traditional publication outlets such as academic journals or edited books can be published.

In 2022 five papers were published, including (a) a paper with one of the keynotes from 2022, (b) a paper that provides an overview of the second year's network activities, and (c) three publications in the series of interviews with web archives about web archives and the COVID-19 pandemic — this unique collection of empirical material now comprises 13 interviews.

In 2022 it was decided to supplement the WARCnet Papers with a parallel publication outlet, WARCnet Special Reports. WARCnet Special Reports were introduced with a view to accommodating for longer reports and similar, and they can be distinguished from the WARCnet Papers by a slightly different cover layout. In 2022 a 212 pages long WARCnet Special Report was published, *Skills, Tools, and Knowledge Ecologies in Web Archive Research*.

Download figures for 2022 indicate that there is a steady interest in the topics covered by WARCnet Papers as well as by WARCnet Special Reports. The five

WARCnet Papers published during 2022 have been visited 24 times on average in 2022 (the average number of visitors for WARCnet Papers published in 2021 was 22), and the WARCnet Special Report was visited 58 times. The number of visits of all WARCnet Papers published since 2020 (Figure 4) indicate that older publications appear to remain relevant to readers. In total, since 2020 all published WARCnet Papers, and WARCnet Special Reports have been visited 938 times.
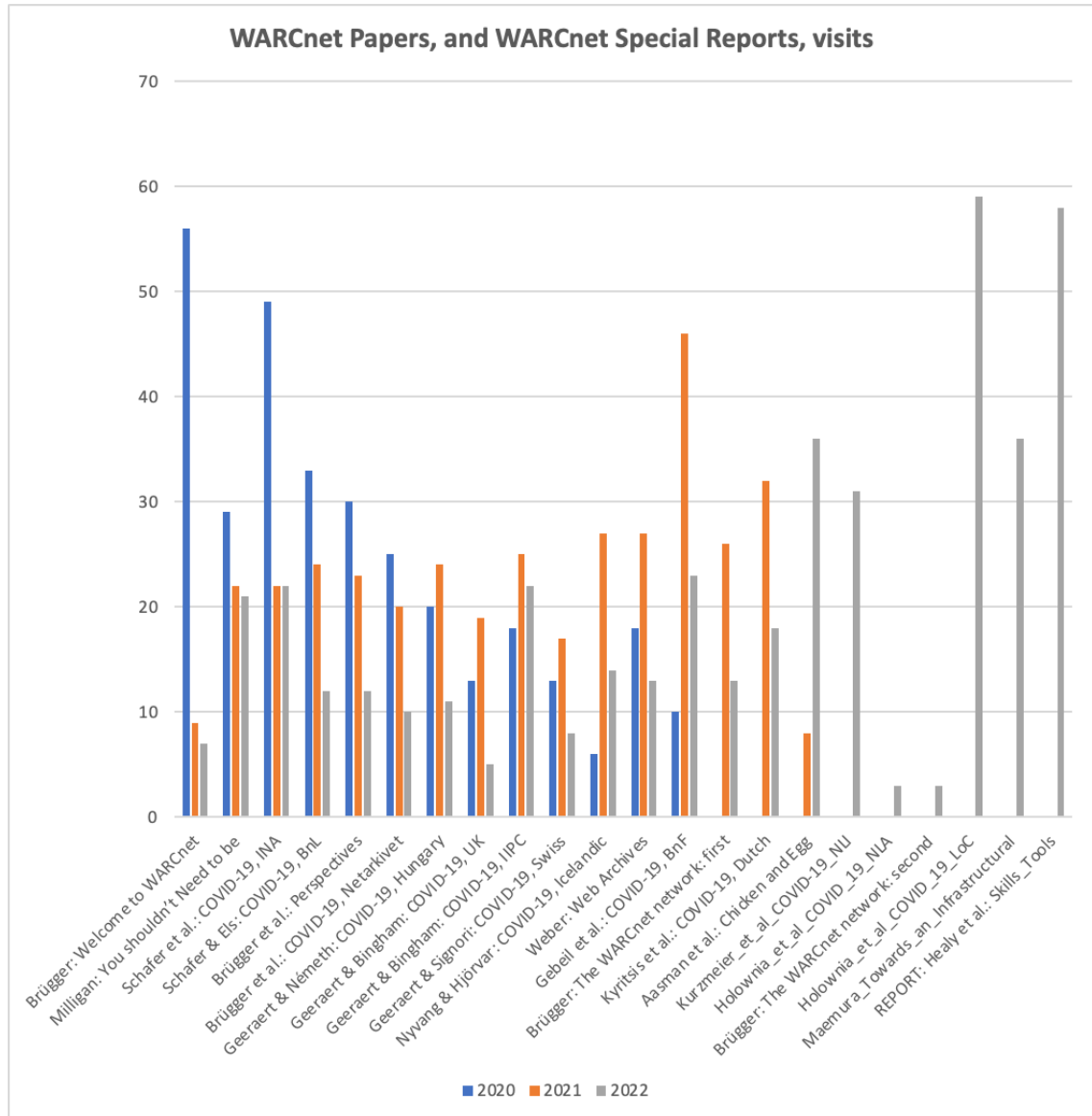

Figure 4: WARCnet Papers, and WARCnet Special Reports, number of visits on warcnet.eu, 2020-2022 (publications in chronological order, from left to right).

Conference papers and publications

In 2022 network members initiated and subitted a number of conference papers, panels, and publications. Some of these publications are initiated by WGs, whereas others are initiated by network members themselves, but are based on work carried out within the network.

Outreach: WARCnet Twitter Challenge

The 'WARCnet Twitter Challenge' was launched in relation to the network meeting in November 2020, and as mentioned above it continued in 2021 and 2022 at network meetings and at the closing conference where the winner of the winners was elected. All submitted entries, nominees, and more related information can be found on twitter.com, #WarcnetChallenge, or on the project website at https://cc.au.dk/en/warcnet/warcnet-twitter-challenge.
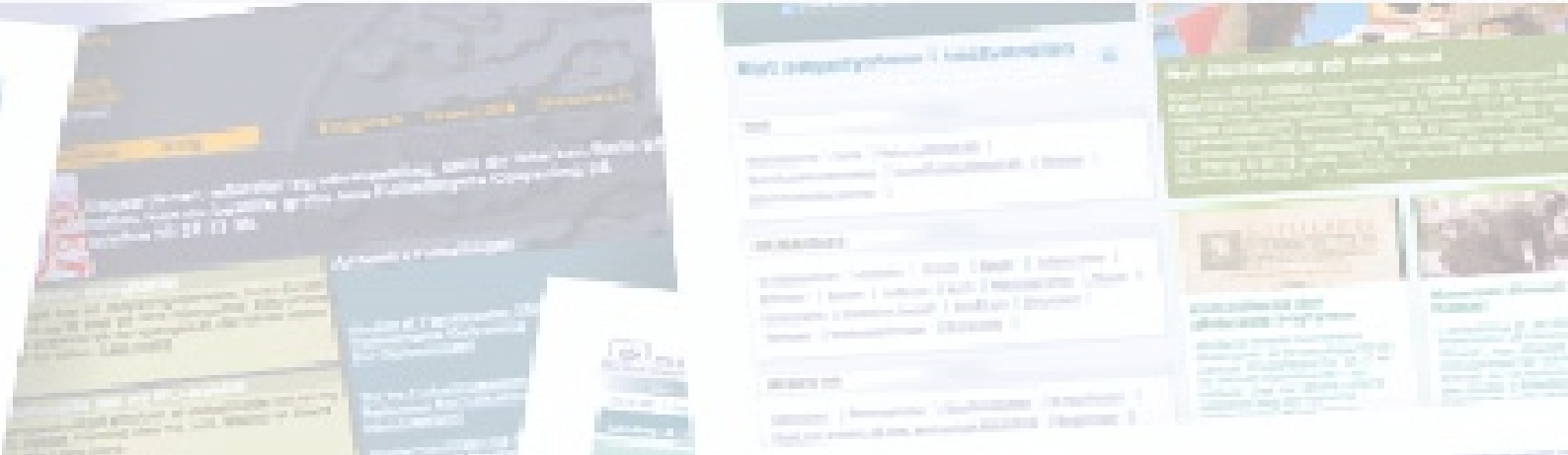
MANAGEMENT

The WARCnet network is headed by three PIs, and together with three network members (one Early Career Researcher, one web archive representative, and one IT developer) they form the Steering Group. The purpose of the Steering Group is to oversee the management of the WARCnet network, including, among others, defining, planning and approving network activities, establishing and implementing a procedure for the addition of new members to the network, and monitoring the project budget. In 2022 the Steering Group has met twice, supplemented with email correcpondance.

A Coordination Group that shall ensure the coordination between the WGs was previously formed, consisting of all Steering Group members and the six WG chairs, and in 2022 it has met twice. Finally, the Short-Term Network Stay Committee that shall manage the allocation of funded network stays has evaluated applications for Short-Term Network Stays on a rolling basis. Last but not least the administration and management of the project is supported by two student helpers, both from Aarhus University.
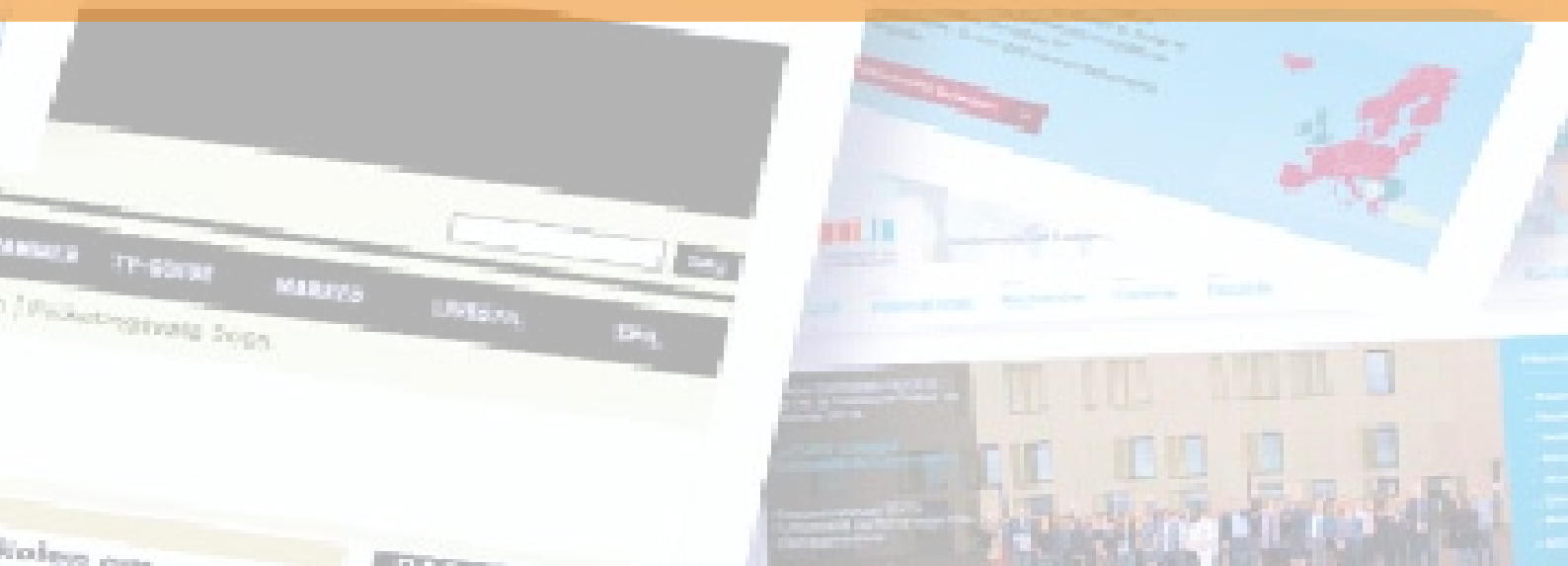
COMING YEAR

As mentioned above the WARCnet network has been able to extend its activities by more than one and a half year. In 2023 the main types of activities are the drafting of book chapters for *The Routledge Companion to Transnational Web Archive Studies*, along with Short-Term Network Stays. In relation to the drafting of book chapters authors will be gathering in Marseille in June 2023, in conjunction with the 5th RESAW Conference that is taking place in Marseille.

Finally, 2023 will hopefully be the year when the WARCnet network is definitely getting out of the shadow of COVID-19.

WARCNET PAPERS

INDEPENDENT
RESEARCH FUND
DENMARK